**Tableau**

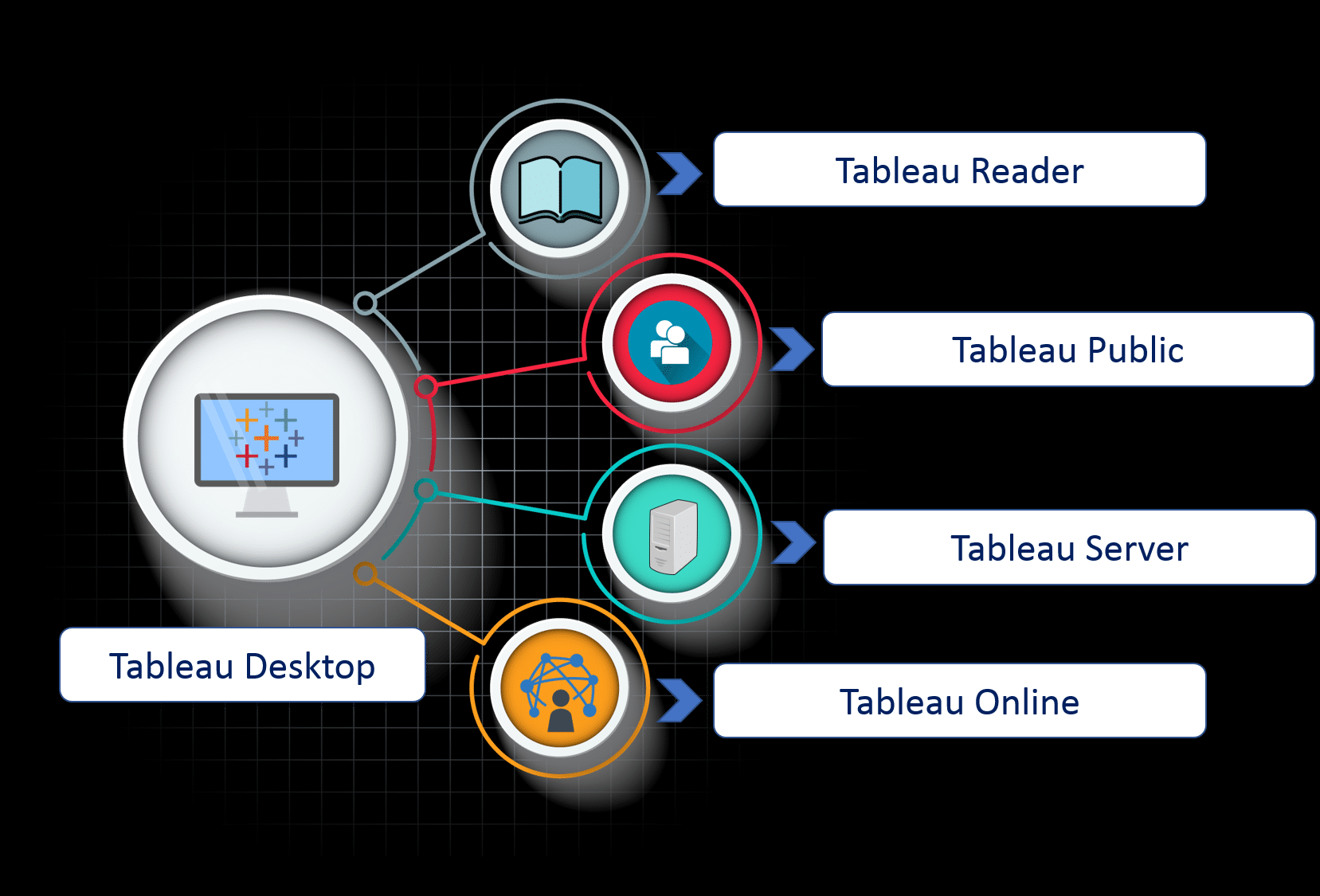
[Data Visualization](https://courses.analyticsvidhya.com/courses/tableau-visual-best-practices-go-from-good-to-great?utm_source=blog&utm_medium=learning-path-tableau-expert) is an art of presenting the data in a manner that even a non-analyst can understand it. A perfect blend of aesthetic elements like colors, dimensions, labels can create visual masterpieces, hence revealing surprising business insights which in turn helps businesses to make informed decisions.

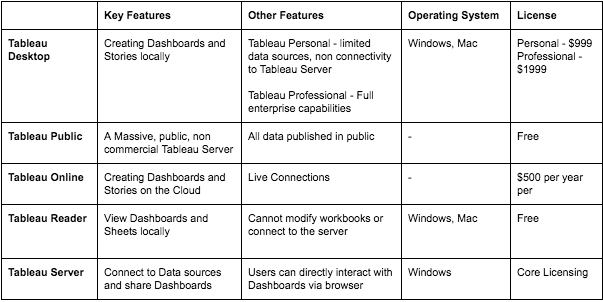
**What is Tableau?**

* Tableau helps in getting insights of data with visualization, helping achieve it from connecting data from number of sources be it flat files to data warehouse.
* **Tableau** can help anyone see and understand their data. Connect to almost any database, drag and drop to create visualizations, and share with a click.

Tableau product family

- Tableau is mostly programmed in C++





**1. Installation**

Out of the five above mentioned products, Tableau Desktop, Public and Online offer Data Visual Creation.

**Tableau Desktop**

It is available in the following three formats :

1. [Free trial for 14 days](https://www.tableau.com/products/trial)
2. If you are a student or a teacher, you get free access to the[Desktop](https://www.tableau.com/academic)for a full year.
3. [Purchase Tableau](https://buy.tableau.com/)

**Tableau Public**

[Tableau Public](https://public.tableau.com/s/download) is purely free of all costs and does not require any licence. But it comes with a limitation that all of your data and workbooks are made public to all Tableau users.

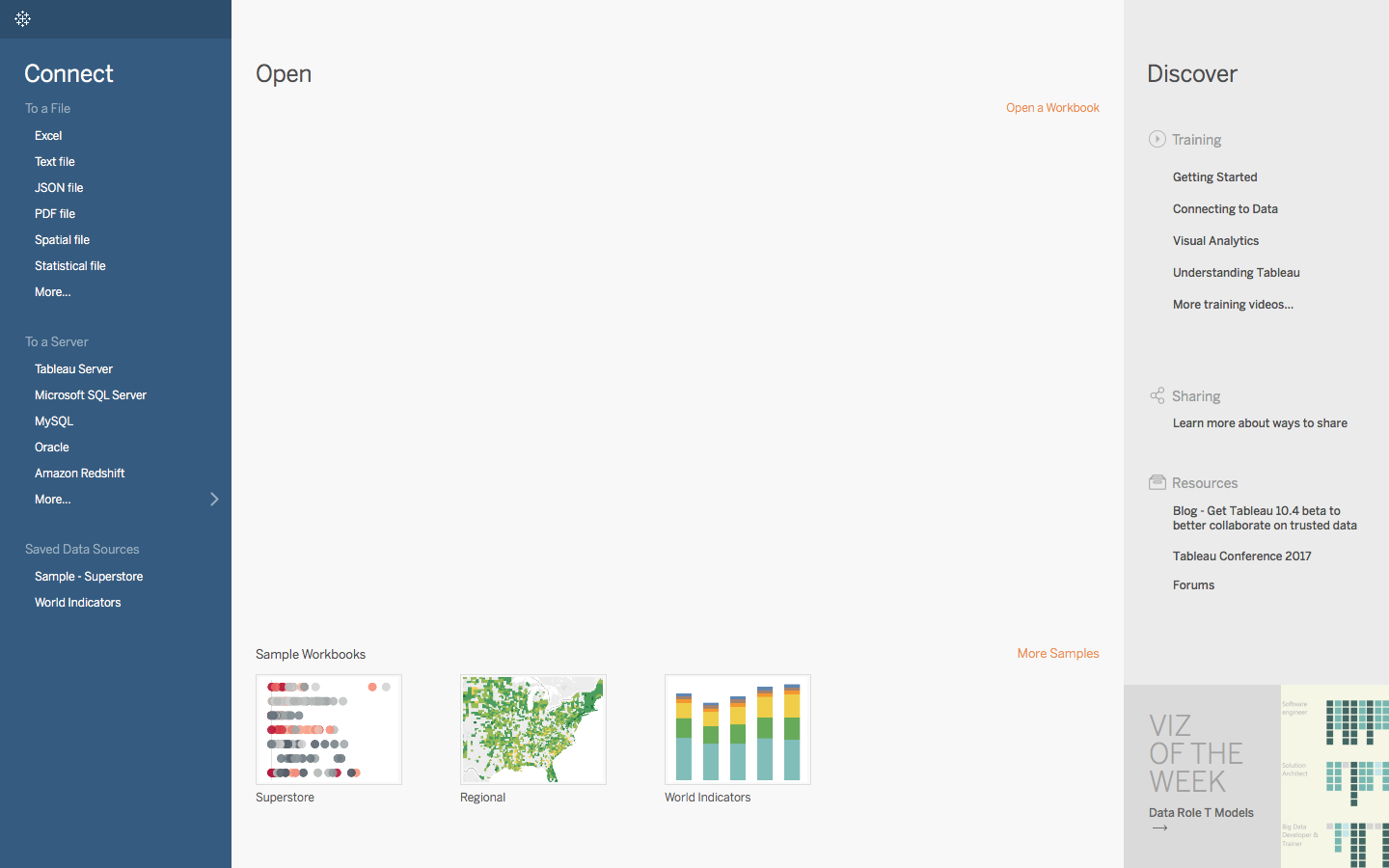
**Tableau Online**

[Tableau Online](https://sso.online.tableau.com/public/idp/SSO) is the best option for you, if you wish to make your Workbooks on the Cloud and be able to access them from anywhere.

**2. Connecting to Data**

Tableau can connect with various data sources such as text, excel file, databases to big data queries also. In this section, we will look at the basics and advance feature of data connectivity with different sources. Here we will also look at Join types, Data Blending, connection with cubes, custom sql and Google Analytics.

When you open Tableau, you will see a screen that looks like this, where you have the option to choose your data connection:

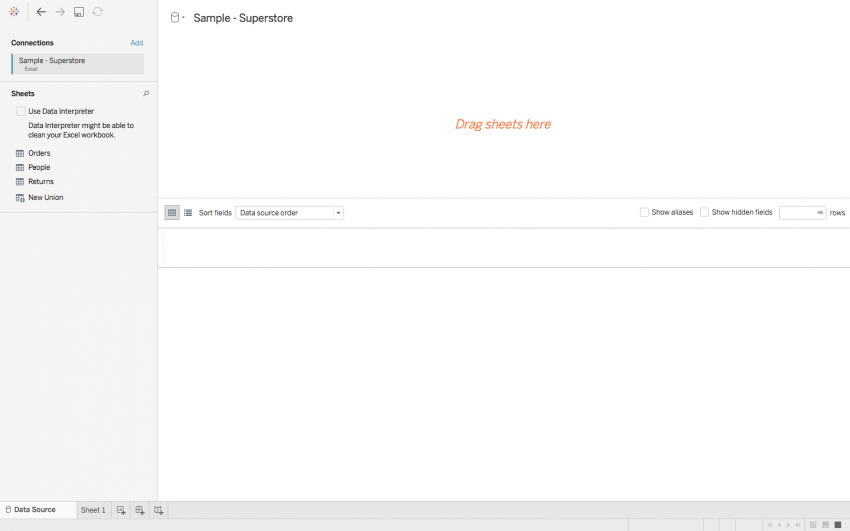


This is where you import your data. As is visible, there are multiple formats that your data can be in. It can be in a flat file such as Excel, CSV or you can directly load it from data servers too.

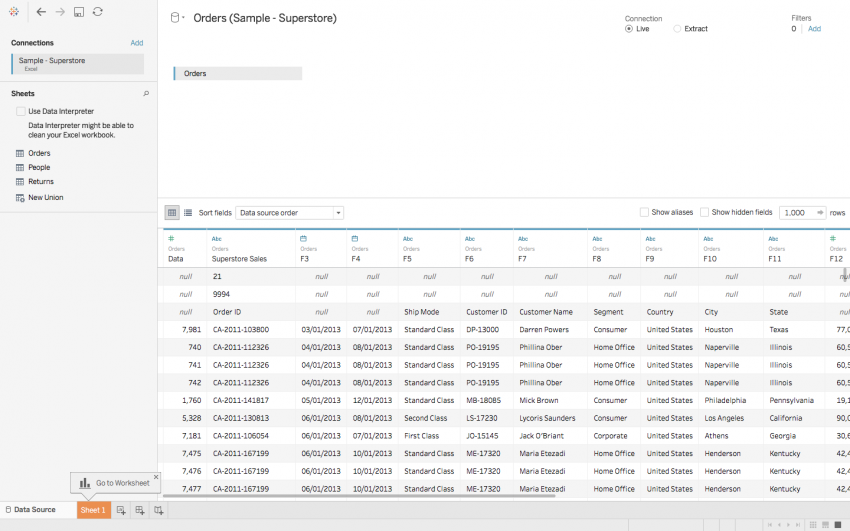
Let’s start with using a sample data set of a superstore, which can be found [here](https://github.com/glued2pc/Tableau-Tutorial/blob/master/Sample-Superstore.xls)

The data is that of a United States’ Superstore which is deliberating over its expansion. It wishes to know the prospective regions of the country. Let’s see how can we visualize the same.

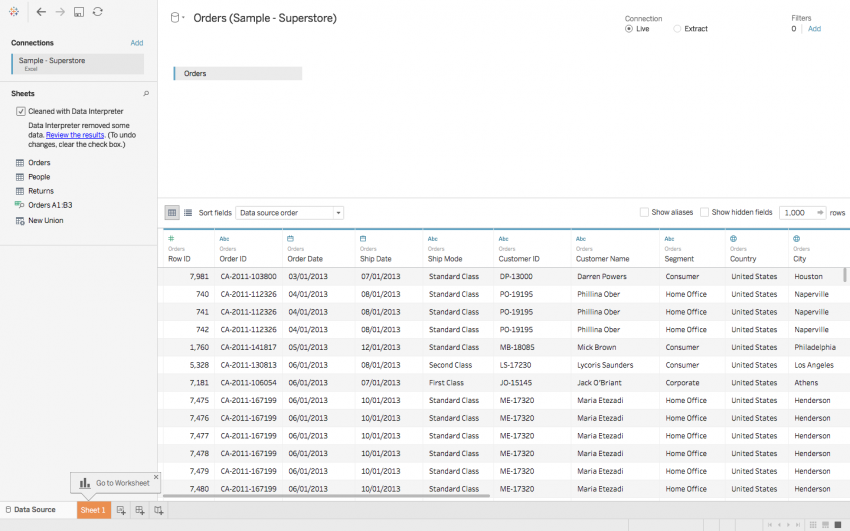
2.1. Import the data, as it is an excel file, click on Connect to a file and choose Excel, once imported the data source Page opens up listing the tables on the left and you see the table data in the right.



2.2. Now drag among the sheet to work on, so let’s pull Orders here.

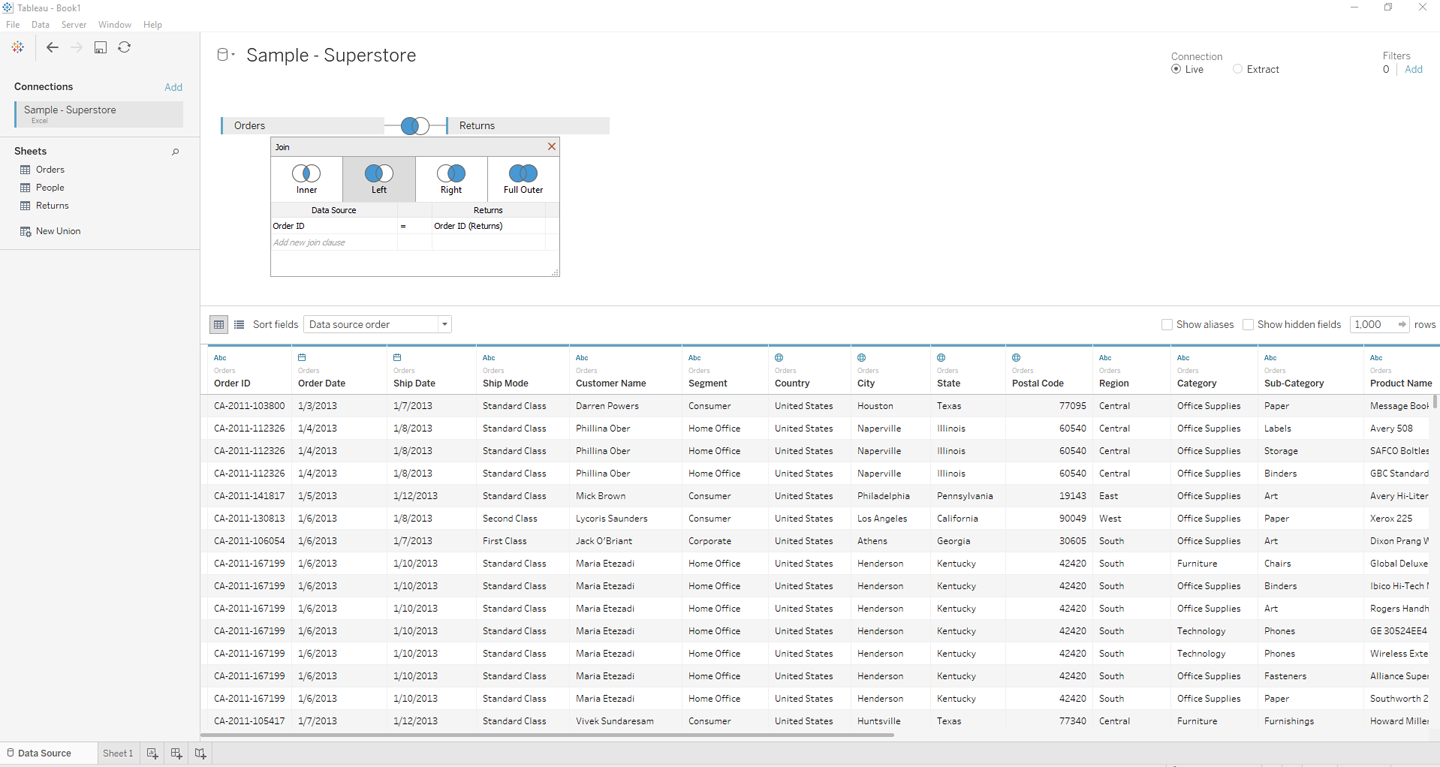


We see there are discrepancies and unwanted rows in the table, to clean the data, we can use the option of **Data Interpreter**, which allows tableau to clean the data automatically.

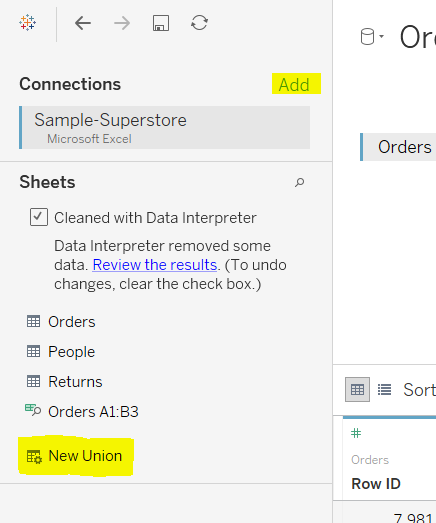


Here we go! Data looks all aligned now.

In Tableau, the Excel workbook is treated as a database and the individual tabs are treated as individual tables within that database. For this reason, you can join tabs to each other if they have at least one field in common. When you join tables, you are appending additional fields to your data source based on shared fields. To do so, simply drag the table (i.e., tab) that you want to join into the data editing interface and tell Tableau what the two tabs have in common.

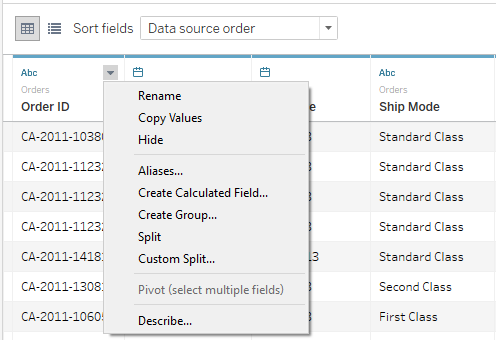


So, here we have left join Orders with Returns sheet, to fetch the matching returns with all the orders, the one which are not available are marked as null.

**Add**: Can be used to add more source data connections and join the

**New Union:** Unions can be used when we have multiple sheets and we know that all column headers are same, instead of joining we can use Union. A new column is added that tells what sheet the data came from.

After retrieving the data, we have option for each column which can be used for Data preparation.

**Rename**: Can change the column name.

**Copy Values**: To copy the initial values in the consecutive rows.

**Hide**: Hide the entire column

**Aliases**: New names to each distinct row value.

**Create Calculated field**: Allows to create a new field with the respective column.

**Split and Custom split**: eg: Name into Firstname and lastname

\*\* Also you can change the Data type of a column by clicking on data type icon, but with caution, because if the data in columns does not match the type, then results in null values.

\*\* and the Changes made to the data at this stage only change the metadata and has no impact on the underlying data source.

**Types of connecting the data**

We have two types of connections available Live or Extract.



**Live** connections are advisable when we are working with large datasets and/or datasets held on powerful, in-memory databases.

**Extracts** create a snapshot of data at the point they are created. Only a part of the whole data is used and so provides a faster approach compared to live connections. Extracts allow us to work when offline and doesn’t require authentication for working on the data each time. Whereas, the data should be refreshed periodically, so the update data is available. This can be refresh tasks can be automated with Tableau server/Online.

**Filters**

This Filters option next to the connection type, is a data source filter. It helps in filtering out the data you don’t need for you analysis. Also, helps in ensuring the security and visibility of data to the required groups only.