**Aim:** Connecting and extracting with various data resources in tableau.

# **Background Information:**

#### **Data Sources**

Tableau can extract data from all the popular data sources. These include:

### 1) File System

The simplest data source you can use with Tableau is a file. These could be files like an Excel spreadsheet, a CSV file or a text file.

### 2) Cloud System

You can also source data from popular cloud sources. Some of the options are:

- Google Analytics
- Google BigQuery
- Windows Azure
- Amazon Redshift

### 3) Relational systems

You can connect to many types of relational databases such as SQL Server, Oracle, and DB2.

#### 4) Live Data Sources

Connect live is a feature of Tableau that allows you to connect real-time data. Tableau does this by constantly reading the data, so your visualizations are constantly up to date.

## 5) Using In-Memory Data

The alternative to connecting to a live data source is to load one into memory. This is a better option for static data that won't change anytime soon, as it will only be loaded once. The in-memory database will then be analyzed by Tableau.

#### 6) Connecting Multiple Data Sources

One of the great features of Tableau is the ability to combine data sources. You can work with data from a file system and data from a relational database all at the same time. All you need to do is define multiple data connections.

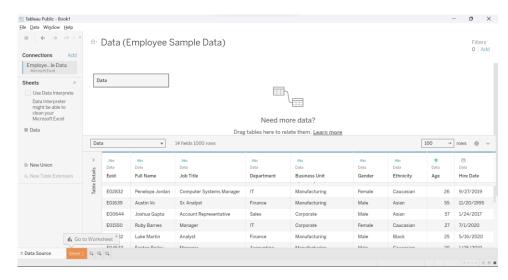
### **Data Extraction Techniques**

- Once you've decided on your data sources, the next step is to extract the data you need from those sources.
- Whether you are connecting to a live database or storing your data in memory, you may well want to cut it down to only what you need for your application. This will mean you'll have less data to extract from a live source or a smaller amount of data to store in memory.
- It also converts the data to a form that works well with the Tableau engine, meaning things will speed up even more.
- With Tableau, this is done with data extracts.
- A data extract is simply a subset of a total data source. When extracting data, you can choose exactly what you want and how much of underlying data to extract using extract data dialog box.
- To create a new Tableau data extract, go to Data -> Extract Data. You'll be presented with many options to limit the number of rows and aggregate for dimensions. Here is where you can use filters to cut down your data to just the things you need.

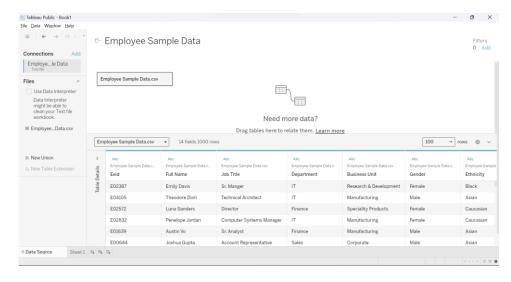
#### **Filtering Extracted Data**

- You might not need every single field and row in the data you've extracted. By cutting it down to just the things you need, you can improve performance and make life easier for yourself.
- There are three main types of filters to use in Tableau
  - 1) Dimension filer
  - 2) Measure filter
  - 3) Date filter
- Each works on a different type of data field. To apply a filter, simply drag a field into the filter pane, it looks like this.
- Then you'll be prompted with some options for your filter. Choose the ones you need and click apply.
- Once you've created a data extract, you can add more data to it from the data pane. Do this by going to Data -> Extract -> Append to File. You can do this with new data types, just make sure they are the same type and have the same number of fields as the original data.
- It's possible to work on large data sets using Tableau. Things do, however, get a little more difficult if your dataset doesn't fit in memory. This is where data extracts and filters really come in handy. If your data is still too big to fit in RAM after extracting and filtering it down, it will still work but will run a lot more slowly.

#### Excel File Data Source



#### Text File Data Source



## JSON File Data Source

