**Tableau Introduction-Assignment 1**

1. **“twbx” is a bundled workbook for Tableau. The original .twb file is bundled with the data source in this package. It can be compared to a compressed file. It contains all the information and instructions required to operate in Tableau. Since the data is contained within the .twbx file itself, one can still access and use it without a network or Internet connection. The .twb file and the data source can be separated from the .twbx file by unpacking it. Pick up any dataset of your choice, create a simple bar chart using the fields of the dataset and save the visualization created in .twbx format. Analyze the properties of the newly created twbx file and segregate the .twbx file into .twb and data source.**

Response:

1. While saving the sheet in Tableau the application actually redirects to the folder

’ C:\Users\User\Documents\My Tableau Repository\Workbooks’. (We can change it, if we wish to ).

1. By saving the sheet in Tableau in .twbx format it got saved with the type-“Tableau Packaged Workbook” and by unpacking in two things saved in the same folder as mentioned above:
   1. Tableau sheet with .twb format and type-“ Tableau Workbook”.
   2. File Folder with the actual data file.
2. **Briefly explain the utility of the Tableau bookmark feature and create a simple bookmark file. Observe the format of the bookmark file and mention the location in which it is saved.**

Response:

1. Tableau bookmark files have the .tbm file extension.
2. By default it gets saved in location-“C:\Users\User\Documents\My Tableau Repository\Bookmarks”
3. While we save a bookmark it saves the snapshot of the worksheet.
4. A bookmark can be opened in a worksheet we are working unlike .twb,.twbx formats opening into a new window.
5. **Using the “Sample-Superstore.xls” file, create a scatter and a bubble plot between different measures in the dataset and observe the type of correlation (negative or positive or no correlation) between them. Draw a comparison between the bubble chart and the scatter plot.**

Response:

1. From the given data source “Sample-Superstore.xls” I have used the sub-category, quantity, discount measures to plot both scatter and bubble plots.
2. The correlation is positive with quantity and discount.
3. Identifying the relationship between the measures quantity and discount at a glance is same in both the charts. But identifying the sub-category is difficult even though the labeling is done as some data points are overlapping in scatterplot.
4. Bubble charts can be customized to make a visually appealing chart by playing with colours.
5. Similarly, scatter plots can be customized by using different shapes. (Even can be added externally).
6. **Consider that you are an HR representative for a multinational company. The staff database is under your control. There are certain details regarding employees that you must never divulge. However, there are many bits of information about employee abilities and skills that may be shared. Using the data extract option in tableau, build a packaged worksheet and use the option "Hide All Unused Fields" in the data extract feature to hide all the fields- dimensions and measures which you haven’t used in the visualization and do not wish to share with employees. Feel free to use any HR dataset or you may even create a dummy data for illustration purpose.**

Response:

1. By following the above steps, able to create a data extract.
2. Tableau Extract type file with ‘.hyper’ extension has been saved in “C:\Users\User\Documents\My Tableau Repository\Datasources” path.
3. Data used from Kaggle.
4. Discuss the differences between the “Measure Names” and “Measure Values” pre-defined features in Tableau. Using the “PowerStore\_USA” dataset available in your iNeuron resources, create a visualization using “Measure Names” and “Measure Values” and mention the fields that fall into each category- “Measure Names” and “Measure Values”.

Response:

1. “Measure Names” and “Measure Values” are both created by Tableau.
2. The difference between the two fields is that Measure Values field generates and displays the values of all the measures and Measure names is a dimension which includes names of all the measures.
3. Measure Names:

* Sales Details
* Customer Name Index
* Quantity
* Unit Cost
* Unit Price

Measure Values:

* Sales Details
* Customer Name Index
* Quantity
* Unit Cost
* Unit Price