ESSENTIALS OF DATA WAREHOUSING AND DATA MINING

LAB ASSIGNMENT - 1

NAME - Kaparotu Venkata Surya Tharani
USN - 22BTRAD018
BRANCH - AI & DE

1) What is tableau ?

ANS - Tableau is a top Business Intelligence (BI) and data visualization tool that can be used by users of all skill levels.
It gives people and organizations the ability to turn unprocessed data into interactive, shared dashboards that offer insights that promote wise decision-making. Tableau prioritizes user-friendliness, allowing both technical and non-technical users to create complex visualizations and analyses with ease.
Using a range of visualization choices, Tableau users can explore trends, patterns, and outliers in their data.

Features of tableau are :
1.Sophisticated data visualization
2. Real time analysis
3. Data Blending

2) What are the different ways of loading data in tableau? Write the steps.

ANS - Different ways of loading data in tableau are :

1. Importing from an Excel Sheet.
2. Importing from a JSON file.
3. Importing from a Text file.

1. Importing from an Excel sheet -

Steps are :

1. To import data from an excel file, first, click on "Microsoft Excel" under the connect tab.

2. It will open a dialog box where you can navigate to the Excel file in your machine from which you want to import the data. Then, click on the file and click on ‘Open’.

3. If there are multiple sheets in your Excel workbook, all of them will be imported automatically and they will be listed as sheets on the left-hand side panel of Tableau.

4. Tableau queries the file for data and imports the data only. So, if any data is updated in the source file, the connection needs to be refreshed each time the data is modified to keep the data up-to-date in Tableau.

5. To add more data (which can be from any source) into the existing workbook, just click on "add" which is present beside connections on the left pane.

6. Sometimes Tableau recommends using the inbuilt data interpreter to use while importing the data. There is no need to worry about it; it justs formats the data.

2. Importing from JSON files. -

Steps are :

1. Importing data from JSON file repeats the first two steps of other formats.

2. First, click on "JSON File" under connect, navigate to the file's location, then select the file, and click “open”. Tableau will prompt you to select the schema levels you want to analyze, so select the levels required and click "ok".

3. There are several differences between importing data from a JSON file and other formats, particularly when dealing with big JSON files.

4. Thus, upon importing a JSON file, Tableau begins by scanning the first 10,000 rows, using these lines to identify the schema.

5. Therefore, we have the option to scan the full file, which may take some time depending on the file size, if you find any fields missing from the detected schema.

6. Tableau will take longer to load the data the more levels there are in the JSON file's format.

3. Importing from Text files. -

Steps are :

1. The first few steps are similar to that of importing data from Microsoft Excel.

2. After clicking on "Text File", navigate to the file that you’re importing, click on "Open", and the data in the text file will be imported. The most frequently used file formats are .txt, .csv, .tab, .tsv.

3. While importing from a text file, the most important property is the delimiter. To change the properties, first, click on the “options” in the text file which will give a drop down and then select the "Text File Properties".

4. From the menu, you can edit the Field Separator, Text Qualifier, Character Set, and Locale. Field Separator is our delimiter.

5. Character Set is the encoding used to encode the text file. In most of the cases, Tableau auto-detects these values.

3) In tableau what is the functionality of "User Data Interpreter" option in 'data source' tab or 'edit data source' option 'Data' dropdown menu.

Ans -

- ➔ Data Interpreter can give you a head start when cleaning your data. It can detect things like titles, notes, footers, empty cells, and so on and bypass them to identify the actual fields and values in your data set.

- ➔ This functionality is available in the "Data Source" tab when you connect to a data source or in the "Data" dropdown menu within the "Edit Data Source" option.

- ➔ The 'User Data Interpreter" option is a data cleaning feature that helps identify and clean up non-data sections in your data source.

- ➔ It can even detect additional tables and sub-tables so that you can work with a subset of your data independently of the other data.

- ➔ The "Data Interpreter" automatically identifies and addresses certain common data preparation challenges such as totals, handling subtotals that might be present in your data source.