

Experiment No.

SHRIDEVI INSTITUTE OF ENGINEERING AND TECHNOLOGY, TUMKUR

Introduction to various Data visualization Tools

Data visualization is the practice of translating information into a visual context, such as map @ graph, to make data easier for the human brain to understand and pull insights from. It is the representation of information and data through use of common graphics, such as charts, plots, infographics, and animations. Data visualization is a powerful way for people, especially data professionals, to display data so that it can be interpreted easily.

Benefits of Data Visualization :

- 1) It is easy to understand the information with graphics
- 2) It made data to be represented in attractive way.
- 3) Shows complex relationships.
- 4) Helps to process large data sets
- 5) Useful for identifying trends

6) Minimizes ambiguity

Data visualization tools provide the ability to see and understand data trends, outliers and patterns in an easy, intuitive way. There are various data visualization tools available. One must choose the tool based on various factors such as its ease of use, types of graphical representations the tool can produce, size of the data set the tool can handle -etc. Some of data visualization tools are Tableau, power BI, Google charts, Jupyter, Graphana -etc.

The following are some common types of data visualization:

Table: A table of data displayed in rows and columns, which can be easily created in a word doc @ excel spreadsheet

Chart @ Graph: Information is presented in tabular form with displayed along x and y axis, usually with bars, points @ lines, to represent data in comparison.

Geospatial visualization: Data is depicted in map form with shapes and colours that

Illustrate the relationship b/w specific locations, such as choropleth @ heat map.

Dashboard: Data and visualizations are displayed, usually for business purposes, to help analysts understand and present data.

Introduction to Tableau and Installation:

Tableau is a data visualization tool that provides pictorial and graphical representations of data. It is used for data analytics and business intelligence.

Tableau provides limitless data exploration without interrupting flow of analysis. With an intuitive drag and drop interface, users can uncover hidden insights in data and make smarter decisions faster.

Tableau is a Business Intelligence tool for visually analyzing the data. Users can create and distribute an interactive and shareable dashboard, which depict the trends, variations and density of the data in the form of graphs and charts. Tableau can connect to flat, relational and Big Data sources to acquire and process data. The S/w allows data blending and real-time

collaboration, which makes it very unique. It is used by business, academic researchers, and many govt organizations for visual data analysis. It also positioned as leader Business Intelligence and Analytics platform in Gartner Magic Quadrant.

Tableau Features:

Speed of Analysis :- As it does not require high level of programming expertise, any user with access to data can start using it to derive value from data.

Self - Reliant :- Tableau does not need a complex s/w set up. The desktop version which is used by most users is easily installed and contains all the features needed to start and complete data analysis.

Visual Discovery :- The user explores and analyses the data by using visual tools like colors, trend lines, charts, and graphs. There is very little script to be written as nearly everything is done by drag and drop.

Blend Diverse Data sets :- Tableau allows you to blend different relational, semi structured and raw data sources in real time, without expensive

up-front integration costs. The users don't need to know the details of how data is stored.

Architecture Agnostic:- Tableau works in all kind of devices where data flows. Hence, the user need not worry about specific h/w or s/w requirements to use Tableau.

Real-Time Collaboration:- Tableau can filter, sort and discuss data on the fly and embed a live dashboard in portals like SharePoint or Salesforce.

Centralized Data:- Tableau server provides a centralized location to manage all of the organization's published data sources. You can delete, change permissions, add tags, and manage schedules in one convenient location. It's easy to schedule extract refreshes and manage them in the data servers.

program 1: Getting started - Tableau workspace, Tableau terminologies, basic functionalities.

Dataset used: vgsales.csv

- 1) Goto start page: Toggle b/w the active sheet and desktop start page
- 2) Data pane: Includes dimensions and measures, populated from your selected data source - may also include calculated fields, parameters (② sets).
- 3) Analytics pane: Includes options you can use to apply reference lines, forecasts, trendlines, etc to add totals to crosstabs, and to build boxplots.
- 4) Workbook names: The file name of our workbook.
- 5) View cards: used for modifying the worksheet
- 6) Toolbar Icons: Icons are available for quick access to popular features.
- 7) Worksheet/ view: workspace for building your visualizations.
- 8) Goto data source: Returns you to data source specification space.
- 9) Worksheet tabs: click to view a specific worksheet, dashboard, (② story).
- 10) New worksheet, dashboard and story tabs: Click to create a new worksheet, Dashboard (② story).

11) Status Bar: Display data about the fields and marks included in the view.

Steps:

1] Tableau workspace setup:

Connect to Data:

- * open tableau and on the "Start page", select connect → To a File → Text File.
- * Browse to the location of vgsale1.csv and open it.

Data preview:

- * After loading, Tableau will show a preview of data. You can rename columns if necessary.
- * Click ~~sheet~~ on the "Sheet 1" tab at the bottom to go to your first worksheet.

2] Tableau Terminologies :

- * Dimensions: These are qualitative fields. In vgsale1.csv, Eg: includes Platform, Genre and publisher.
- * Measures: These are quantitative fields used for calculations, Eg are Global-Sale, NA-Sale, and Year
- * Rows and columns shelf: Drag dimensions and measures to the Rows @ columns shelves to build the structure of your visualization.

* Marks: Controls the appearance of the data. You can set marks to be circles, bars or other shapes and control size, color and label.

* Filters: Used to filter the data displayed in the view.

* Pages shelf: used for creating animations or segmenting your view by categories.

3] Basic Functionalities:

a) Basic visualization (Bar chart of Global Sales by genre):

- * In your worksheet, drag Genre to the Column shelf.
- * Drag Global_Sales to the Rows shelf.
- * You should see a bar chart. If the data isn't aggregating correctly, check if the aggregation is set to sum by right-clicking Global_Sales → Measure → Sum.

b) Sorting:

- * Click on the Global_Sales axis and sort descending to show the genres with the most sales first.

c) Filtering:

- * Drag Year to the Filter shelf.
- * Choose the range of years you want to display (Eg: 2000-2010)
- * Add Year to the Page shelf to create a dynamic

view of how sales changed over time.

4] Additional Functionality:

Dashboards: Combine different sheets to create a comprehensive dashboard. Goto the Dashboard tab, drag your created sheet to the layout, and arrange them accordingly.

a) Add one more worksheet- Global Sales Trend by year

- * Drag year to the Columns.
- * Drag Global-Sales to the Rows.
- * Create a line chart to how global sales have trended over time.
- * Add Generes to the Marks.
- * Apply color to Generes.

b) Goto the Dashboard tab in Tableau.

c) Add multiple visualizations to a single dashboard.

d) Arrange charts.

Program 2: Connecting to Data Source - Connecting to Data base,
Different types of Tableau joins.

Dataset used: Tableau Join file: Contain 3 sheets: Demogra
-phice, salary, job title.

1] Connecting to Excel File in Tableau :

- * Open Tableau and click on Connect in the left pane.
- * Under to a file, choose Microsoft excel.
- * Browse and select your excel file
- * Tableau will display the sheets from the Excel file in the Data source tab.
- * Drag the relevant sheet to the workspace.

2] Tableau Joins file.xlsx Data set : having three Excel sheets.

Demographics:

- * Employee ID
- * Name of Employee
- * Employee Age
- * Employee Gender.

Salary:

- * Employee ID
- * Employee Salary.

These sheets have a relationship based on the EmployeeID, and can join them using this field.

Now Drag and drop salary table - that allows you to do join of your choice.

3) Types of joins in Tableau:

Once both tables are in the Data Source tab, Tableau automatically suggests an inner join, but can you modify the type of join depending on the scenario.

a) Inner Join :

* Description: Returns only records where there is a match in both ~~data~~ tables.

* How to create in Tableau:

* Drag Demographics and Salary sheets into the canvas.
* Tableau automatically detects the common field. If not, manually select it.

* Choose Inner join in the Join Type options.

* Result: You will see only employees id matches in both Demographics and Salary table

b) Left Join :

* Description: Returns all records from the left table and matched records from right table . If there's no match,

NULL values are returned for fields from the right table

How to create in tableau:

- * In the join settings, select left join.
- * Result: All employees will be returned, even if data missing in Salary. Salary information will be NULL for those without a match.

c) Right Join:-

* Description: Returns all records from the right table, and matched records from the left table. If there's no match, NULL values are returned for fields from the left table

How to create in Tableau:

Select Right Join

Result: You will see all salary, even if they don't have employee id. Employee info. will be NULL for those salary with no matching employee id.

d) Full outer Join:

* Description: Returns all records when there is a match in either left or right table. If there's no match,

NULL values are returned for missing value.

* How to create in Tableau:

- * Select Full outer join

Result: You will see all employees and all salary, even if they don't have a match in the other table, NULL values will appear

4] Creating a visualization Based on ~~the~~ joins:

After performing the joins, you can build different visualizations:

Press on sheet 1:

Eg:

* Bar chart: No. of employees and their salary

* Drag Name of Employee to columns.

* Drag Employee Salary to Rows

* This chart will display the number of employees and their salary based on the types of join.

* Sort it in descending

* Drag Employee Salary to Marks - select color, color, label.

program 3: Creating a view - formatting charts, adding filters, creating calculated fields and defining parameters.

Step 1: Connect to Data

- 1) Open Tableau Desktop
- 2) Connect to your Data Source:
 - a) Click on Connect on the left sidebar
 - b) Choose your data source by selecting text file and load your vgsales dataset into Tableau.

Step 2: Create a Basic visualization

* Create a new worksheet:

- a) Click on the sheet tab at the bottom of the screen.

* Drag Fields to shelves:

- a) Drag year to columns shelf.
- b) Drag Global Sales to the Rows shelf.
- c) Drag EU Sales to the Rows shelf.

That gives the line graph visualization.

* Change visualization type:

① In the show me panel on the right, select a bar chart or any other type that suits your needs.

Step 3: Format the chart

* Format Axis:

- a) Right-click on the Global Sales axis and select Format.
- b) In the Format pane, adjust the font style & size as needed.

* Add Title and Annotations:

- a) Click on the chart title area and enter a descriptive title - Global Sales by Year.
- b) Add annotations if needed to highlight specific data points. - Right click on the chart which you want to highlight - select Annotate - select mark - press ok.

Step 4: Add filters

Add a filter for year:

Drag year to the filters shelf.

Choose the range of years you want to display

Step 5: Create Calculated Fields

*Create a calculated field for Sales category:

- a) Right-click on the Global Sales - Select - Create calculated field
- b) Give name to your calculations as Global Sales - EU Sales
- c) Do calculations as per your need - [Global Sales] - [EU Sales]
- d) Press ok

Add calculated fields to visualization:

- a) Drag Global Sales - EU Sales to the Rows shelf to show Global Sales over year with Global Sales - EU Sales Over 6 Years

Step 6: Create a Parameter:

Name: "Select Genre"

Data Type: String

values: List (e.g: "action", "Adventure", "Shooter") @ Add values from Genre.

Create a calculated field:

Name: "Sales by Genre"

formula: IF [Genre] = [Select Genre] THEN [Global Sales]
ELSE 0 END

Build the visualization:

* Columns: Drag "Year"

* Rows: Drag "Sales by Genre".

* At right side of your sheet you can select required
Genre and can see different visualization.