Data Visualisation

unit: 01

Module Name: Introduction to Data Visualisation

Aim

To familiarise students about the basic terms and concepts of Data Visualisation.

Objectives

The objectives of this module are to understand:
- History of Data Visualisation.
- Scientific design choices.
- Learn Adavnced Excel, learn Power Bi, Learn Tableau

Outcome

At the end of this module, you are expected to:

- Describe the need of Visualisation techniques.

- Explain various graphical forms.

Introduction to Data Visualisation

Content

Unit 1: Introduction to Data visualization using Excel

Introduction to Data Visualization, Story Telling with Data, Pivot Table, Charts and Power query, Data Connection using Excel and Dashboarding

Introduction to Data Visualisation

- Data Visualisation is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data Visualisation tools provide an accessible way to see and understand trends, outliers, and patterns in data.

- Effective data Visualisation should be informative, efficient, appealing, and in some cases interactive and predictive.

Data visualization is a graphical representation of any data or information.

Visual elements such as charts, graphs, and maps are the few data visualization tools that provide the viewers with an easy and accessible way of understanding the represented information.

Data visualization enables you or decision-makers of any enterprise or industry to look into analytical reports and understand concepts that might otherwise be difficult to grasp.

Introduction to Data Visualisation

- In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

- Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns.

- With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed.

- Its actually dates back from the 17th Century

Introduction to Data Visualisation

- **Common Tools:**
- Microsoft Excel
- Power BI
- Tableau
- Python libraries (matplotlib, seaborn)
- **Key Features:**
- Scalability: Works with small and large datasets.
- Interactivity: Filters, slicers, and hover effects enhance user engagement

•

Data visualization is the graphical representation of data. It is used to communicate information to an audience in a clear and concise way Data visualization can be used for a variety of purposes, such as:

- To present data to a large audience
- To help people understand complex data
- To identify trends and patterns in data
- To make decisions based on data

There are many different types of data visualization techniques, each with its own advantages and disadvantages. Some of the most common types of data visualization techniques include:

- Bar charts
- Line charts
- Pie charts
- Scatter plots
- Maps
- Heatmaps

- **Key Components:**

- **Data:** Accurate, relevant, and clean datasets form the foundation.

- **Visualizations:** Select appropriate charts and graphs that highlight key points.

- **Narrative:** Tie visuals to a clear story, using captions or presentations.

- **Context:** Provide the background or frame of reference for the data.

- **Techniques:**

- Use color coding to emphasize key data points.

- Leverage annotations to explain outliers or trends.

- Maintain simplicity to avoid overwhelming the audience.

-
    When choosing a data visualization technique, it is important to consider the following factors:

    - The purpose of the visualization
    - The audience for the visualization
    - The type of data being visualized
    - The available resources

    Data visualization is a powerful tool that can be used to communicate information effectively. By choosing the right technique and using it effectively, you can create visualizations that will help your audience understand and interpret your data.

Introduction to Data Visualisation

Why Excel for Data Visualization?Widely used tool for analytics and reporting.Supports advanced features like Pivot Tables, Power Query, and Dashboards.Easy integration with other tools and data sources.

What is an Excel Pivot Table?

- An interactive worksheet table
  - Provides a powerful tool for summarizing large amounts of tabular data
- Similar to a cross-tabulation table
  - A pivot table classifies numeric data in a list based on other fields in the list
- General purpose:
  - Quickly summarize data from a worksheet or from an external source
  - Calculate totals, averages, counts, etc. based on any numeric fields in your table
  - Generate charts from your pivot tables

pivot table

Excel pivot tables are a feature that you should learn how to use. Instead of analyzing countless spreadsheet records, these tables can aggregate your information and show a new perspective in a few clicks. You can also move columns to rows or vice versa. The problem is people believe creating a pivot table is difficult to learn. *for example worksheet*.)

Introduction to Data Visualisation

- **PivotTables** can help make your worksheets more manageable by **summarizing** your data and allowing you to **manipulate** it in different ways.

- A Pivot Table is way to present information in a report format.

- PivotTable reports can help to **analyze numerical data** and answer questions about it.

- Eg:
  - Who sold the most, and where.
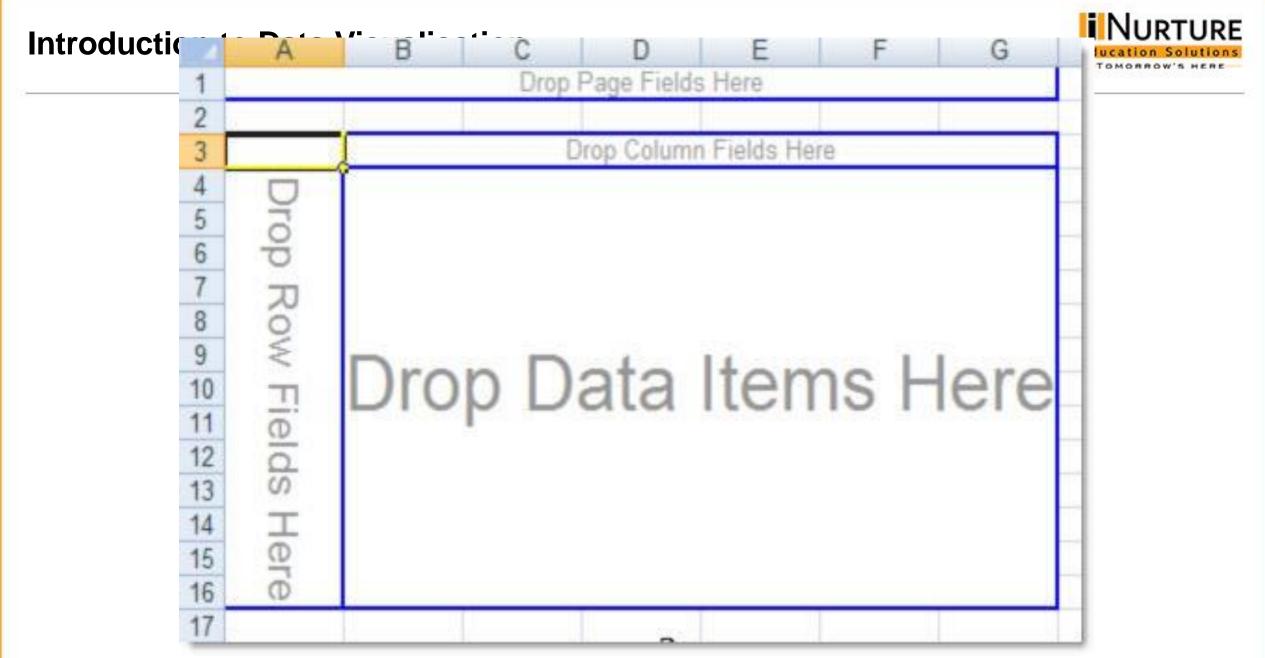 - Which quarters were the most profitable, and which product sold best.

Definition:Pivot tables are interactive tables that summarize large amounts of data quickly, enabling data analysis and reporting.

- feature to summarize, analyze, and explore data interactively.

- Key Features:Summarize large datasets.Drag-and-drop fields for flexible layout.Built-in calculations (sum, average, count, etc.).

- Use Cases:Sales performance by region.Product category analysis.

- Steps to Create:Select your dataset.Go to Insert > PivotTable.Drag fields into rows, columns, values, and filters

Introduction to Data Visualisation

- Pivot Tables in Excel

- Steps to Create a Pivot Table:
- Insert Pivot Table:Select your dataset.Go to the Insert tab > Pivot Table.
- Drag Fields: Drag columns into Rows, Columns, Values, and Filters areas.
- Apply Filters and Sort:
- Use slicers or field settings to refine the view.
- Uses: Summarizing data (e.g., sales by region).
- Calculating metrics like sums, averages, counts. Comparing data points across different categories.
- Example:A dataset of product sales can be summarized to show total sales by category and region.

Where to place data fields

- Page Fields: display data as pages and allows you to filter to a single item

- Row Fields: display data vertically, in rows

- Column Fields: display data horizontally, across columns

- Data Items: numerical data to be summarized

Pivot Table Advantages

- **Interactive**: easily rearrange them by moving, adding, or deleting fields

- **Dynamic**: results are automatically recalculated whenever fields are added or dropped, or whenever categories are hidden or displayed

- **Easy to update**: "refreshable" if the original worksheet data changes

Introduction to Data Visualisation
Pivot Table Geography

	PivotTable Geography
1	The pivot table will be located here. The size will adjust as it needs to automatically. • NOTE: If you move your mouse out of this area, the PivotTable Field List will disappear. To get it back, merely left click your mouse within this area again and it will appear.
2	The fields listed here are your column headers on your original data source. • These fields may be utilized in designing your PivotTable Report. • You may use a field more than once.
3	Report Filters: • Similar to a mentally page break. Allows the user to classify the data
4	Column Labels: • Often created automatically by dragging data fields to the value zone. The user may also drag data fields to this zone for grouping, etc.
5	Row Labels: • Most common label
6	Values: • Wide range of calculations may be performed on the values dragged to this zone.

Sample Data

Pivot Table-Formulas-Calculated Fields

	A	B	C	D	E	F	G	H	I	J	K
1	Policy	Start	Expiry	Location	State	Region	InsuredValue	Construction	BusType	EQ	Flood
2	100682	17-Jul-07	11-May-08	Urban	NY	East	$ 62,550,000	Fire Resist	Apartment	Y	Y
3	100502	7-May-07	31-Mar-08	Urban	NY	East	$ 57,343,200	Fire Resist	Office Bldg	Y	Y
4	100275	2-Feb-07	25-Jan-08	Urban	WI	Midwest	$ 53,410,614	Frame	Construction	Y	Y
5	100625	27-Jun-07	7-Apr-08	Urban	NY	East	$ 49,837,500	Fire Resist	Apartment	Y	Y
6	100342	3-Mar-07	13-Jan-08	Urban	NY	East	$ 39,642,500	Fire Resist	Apartment	Y	Y
7	100574	8-Jun-07	25-Mar-08	Urban	VT	Northeast	$ 37,785,000	Frame	Apartment	Y	Y
8	100754	10-Aug-07	29-May-08	Urban	NY	East	$ 36,909,180	Metal Clad	Manufacturing	Y	Y
9	100498	3-May-07	6-Feb-08	Urban	NJ	East	$ 36,356,000	Fire Resist	Apartment	Y	Y
10	100678	15-Jul-07	21-Apr-08	Urban	WI	Midwest	$ 35,245,000	Masonry	Apartment	Y	Y
11	100596	18-Jun-07	28-Mar-08	Rural	MI	Central	$ 34,841,710	Metal Clad	Manufacturing	Y	Y
12	100214	5-Jan-07	23-Oct-07	Urban	NY	East	$ 33,250,000	Frame	Apartment	Y	Y
13	100428	7-Apr-07	13-Jan-08	Urban	NJ	East	$ 33,000,000	Fire Resist	Service	Y	Y
14	100957	5-Nov-07	5-Nov-08	Urban	NY	East	$ 31,285,000	Frame	Office Bldg	Y	Y
15	100741	5-Aug-07	5-Jun-08	Urban	NY	East	$ 31,265,600	Fire Resist	Apartment	Y	Y
16	100818	2-Sep-07	2-Sep-08	Urban	IL	Midwest	$ 30,128,000	Fire Resist	Apartment	Y	Y

Introduction to Data Visualisation

- Charts visually represent data and make it easier to identify trends, patterns, and relationships.
- **Common Chart Types:**
- **Bar and Column Charts:** Compare quantities across categories.
- **Line Charts:** Show trends over time.
- **Pie Charts:** Represent proportions within a dataset.
- **Scatter Plots:** Show relationships between two variables.
- **Steps to Create a Chart:**
- Select the dataset.
- Go to **Insert** tab > Choose a chart type.
- Customize using the **Chart Tools** for titles, labels, and styles.
- .

- Power Query is a data transformation tool in Excel that helps clean, combine, and prepare data for analysis.

- **Features:**

- **Data Cleaning:** Remove duplicates, fill missing values, and split/merge columns.

- **Data Transformation:** Change data types, filter rows, and unpivot columns.

- **Data Connection:** Connect to various sources like Excel files, databases, and web services.

- **Steps to Use Power Query:**

- Go to the **Data** tab > Click **Get Data** > Select the source.

- Use the Power Query Editor to clean and transform the data.

- Load the processed data back into Excel.

- Excel allows connecting to external data sources to fetch and analyze real-time data.
- **Supported Data Sources:**
- SQL databases
- Web services
- Google Sheets
- Other Excel workbooks
- **Steps to Connect Data:**
- Go to the **Data** tab > Click **Get Data**.
- Select the source (e.g., Web, SQL Server).
- Provide credentials and select the required tables.
- Load the data into Excel for analysis.

Introduction to Data Visualisation

- Dashboards are interactive, visual interfaces that consolidate and display data summaries for decision-making.

- **Steps to Create a Dashboard:**

- **Prepare Data:** Clean and format data for analysis.

- **Create Pivot Tables:** Summarize data for key metrics.

- **Insert Visuals:** Add charts, slicers, and conditional formatting.

- **Design Layout:** Organize visuals and add titles or labels.

- **Add Interactivity:** Use slicers and dynamic ranges to enable filtering.

- .

Introduction to Data Visualisation

- **Best Practices for Dashboarding**

- **Design Tips:**
  - Use consistent colors and fonts.
  - Focus on key metrics and avoid unnecessary details.
  - Ensure interactivity is intuitive.

- **Common Mistakes:**
  - Overloading with too much data.
  - Poor choice of visualizations.
  - Ignoring audience needs.

Introduction to Data Visualisation

- **Case Study Example**
- **Scenario:**
  - Sales performance dashboard for a retail business.
- **Steps Taken:**
  - Data imported using Power Query.
  - Sales summary created with Pivot Tables.
  - Charts added for visual insights.
  - Interactive filters implemented using slicers.
- **Outcome:**
  - Improved decision-making with clear and actionable insights.

Introduction to Data Visualisation

Introduction to Data Visualisation

Assignment

1. Write a short note on Data Visualisation.

2.

Summary

- Data Visualisation is the graphical representation of information and data.

- Effective data Visualisation should be informative, efficient, appealing, and in some cases interactive and predictive.

- Towards the end of 1700–1799, we saw the first attempts at the thematic mapping of geologic, economic and medical data.

- 1850–1900 was the Golden Age of Statistical Graphics.

- From 1975, High-D, Interactive and Dynamic Data Visualisation techniques toll place.

- A grammar of graphics is a tool that enables us to concisely describe the components of a graphic.

Summary

- A good grammar will allow us to gain insight into the composition of complicated graphics, and reveal unexpected connections between seemingly different graphics.

- So many scientific design in data Visualisation were invented.

- There are three types of bar graphs: Horizontal (left to right), Column (up and down) and Stacked (which can be either).

- Scatter plot is not only fun to say, it is what you need when you are looking for the correlation in a large data set.

- A Histogram is a vertical bar chart that depicts the distribution of a set of data.

- A grammar of graphics is a tool that enables us to concisely describe the components of a graphic.

Introduction to Data Visualisation

Document Links

Topics	URL	Notes
Introduction to Data Visualisation	http://www2.cs.uh.edu/~gnawali/courses/cosc6397-f13/intro-visualization.pdf	This web link will describe data Visualisation from very basic.
History of Data Visualisation		This web link will provide entire history of evolution of data Visualisation techniques.
Grammar of graphical techniques	https://byrneslab.net/classes/biol607/readings/wickham_layered-grammar.pdf	This web link will provide details description about grammar of graphical techniques.
Comparison of graph Representation techniques	http://shodhganga.inflibnet.ac.in/bitstream/10603/143688/2/file%202%20chapter%201%20data%20representation%20techniques.pdf	This web link will provide comparison of graph Representation techniques.

Introduction to Data Visualisation

Video Links

Topics	URL	Notes
Introduction to Data Visualisation	https://www.youtube.com/watch?v=XIgjTuDGXYY	This web link will describe data Visualisation from very basic.
History of Data Visualisation	https://www.youtube.com/watch?v=N00g9Q9stBo	This web link will provide entire history of evolution of data Visualisation techniques.
Grammar of graphical techniques	https://www.youtube.com/watch?v=0gGSpJvT4d4	This web link will provide details description about grammar of graphical techniques.
Comparison of graph Representation techniques	https://www.youtube.com/watch?v=Gm1yDbQJL0M	This web link will provide comparison of graph Representation techniques in short.

Introduction to Data Visualisation

E-book Links

Topics	URL	Notes
Data Visualisation	http://haralick.org/DV/Handbook_of_Data_Visualisation.pdf	All Pages
History of Data Visualisation	https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/49286/IDL-49286.pdf	This web link will provide entire history of evolution of data Visualisation techniques.