Jaya Sai Naga Lakshmi katrela

05/03/2024

Data Visualization

Table of Contents

**Introduction To Data**

1. Types of Data
2. Data Structures
3. Data Manipulation

**Introduction To Excel**

1. Importing Data
2. Data Cleaning and Formatting
3. Data Visualization

Introduction To Tableau

1. Connecting to Data Sources
2. Creating Visualizations
3. Interactive Dashboards
4. Sharing and Collaboration

**Policies and Procedure-Task (1)**

**Task-2**

1. Workbook protection
2. Changing currency format
3. Changing sheet to table format and filtering the data
4. Creating charts

**Task-3**

1. Creating Macros
2. Setting header and footer

**Task-4 (Tableau)**

1. Importing Data
2. Creating charts
3. Publishing to tableau public

**Introduction to Data:**

Data refers to facts, observations, or information that can be collected, stored, and analysed. It is essential in various contexts, including business, research, science, and everyday life. Data can come in different forms and formats, such as text, numbers, images, and multimedia. Understanding data is crucial for making informed decisions, identifying patterns, trends, and relationships, and gaining insights into various phenomena.

**Types of Data:**

Data can be classified into different types based on its nature and characteristics:

* **Textual data** can be analyzed to understand sentiment, frequency of terms, and categorization.
* **Numerical data** is used for quantitative analyses, such as statistical modeling, financial forecasting, and measuring performance metrics.
* **Boolean data** helps in filtering, segmentation, and classification tasks, enabling analysts to create more tailored and precise analyses.
* **Date/Time data** is crucial for time series analysis, trend analysis, and scheduling optimizations.

**Data Structures:**

Data structures are fundamental concepts in computer science and data analysis that organize and store data efficiently. Common data structures include:

**Arrays**: Arrays store a collection of elements of the same data type in contiguous memory locations.

**Lists**: Lists are collections of elements that can vary in size and type and can be easily manipulated.

**Stacks**: Stacks follow the Last In, First Out (LIFO) principle and support operations like push (add) and pop (remove).

**Queues**: Queues follow the First In, First Out (FIFO) principle and support operations like enqueue (add) and dequeue (remove).

**Trees**: Trees are hierarchical data structures composed of nodes connected by edges, commonly used for organizing hierarchical data.

**Graphs**: Graphs consist of vertices (nodes) and edges (connections) and are used to represent relationships between objects.

Understanding data structures helps in efficient data storage, retrieval, and manipulation, which is essential for programming and data analysis tasks.

**Data Manipulation:**

Data manipulation involves transforming raw data into a more meaningful format for analysis. Common data manipulation techniques include:

**Sorting**: Rearranging data in a specific order based on certain criteria.

**Searching**: Finding specific data elements based on given conditions or criteria.

**Filtering**: Selecting a subset of data that meets certain criteria or conditions.

**Merging**: Combining multiple datasets into a single dataset based on common attributes or keys.

These techniques are essential for preparing data for analysis, identifying patterns, and extracting insights. Understanding data manipulation helps in cleaning, transforming, and integrating data from different sources for analysis purposes.

**Introduction to Excel:**

Microsoft Excel is a powerful spreadsheet software widely used for data analysis, visualization, and reporting. It provides a user-friendly interface for organizing, manipulating, and analysing data sets. Excel allows users to perform various data-related tasks efficiently, making it a valuable tool for professionals in diverse fields such as finance, business, education, and research.

**Importing Data:**

Excel offers several methods for importing data from external sources:

**From CSV files**: Users can import comma-separated values (CSV) files directly into Excel using the "Data" tab and selecting "From Text/CSV" option.

**From Databases**: Excel supports importing data from various databases such as SQL Server, Access, Oracle, etc., by establishing a connection or using specific import wizards.

**From Web Sources**: Data can be imported from web pages or web services using the "From Web" option under the "Data" tab.

**Data Cleaning and Formatting:**

Data imported into Excel may require cleaning and formatting to ensure accuracy and consistency. Common techniques include:

**Removing Duplicates**: Excel provides built-in tools to identify and remove duplicate values within a dataset.

**Handling Missing Values**: Users can handle missing or incomplete data by deleting rows, filling missing values with appropriate replacements, or using statistical methods.

**Ensuring Consistency**: Techniques such as text-to-columns, data validation, and conditional formatting can be used to ensure consistency in data formats and values.

**Data Analysis**:

Excel offers various features for basic data analysis tasks:

**Sorting**: Users can sort data based on specific criteria to organize it in ascending or descending order.

**Filtering**: Excel provides filter options to display only the data that meets specific criteria, making it easier to analyse subsets of data.

**Performing Calculations**: Users can perform common calculations such as sum, average, count, etc., using built-in functions or formulas.

**Data Visualization:**

Excel includes basic visualization tools to create informative charts and pivot tables:

**Creating Charts**: Users can create various types of charts, including bar charts, pie charts, line graphs, scatter plots, etc., to visualize data trends and patterns.

**Pivot Tables**: Pivot tables allow users to summarize and analyse large datasets by rearranging and summarizing data dynamically.

**Introduction to Tableau**:

Tableau is a powerful data visualization and analytics tool that allows users to create interactive and shareable dashboards, reports, and visualizations from various data sources. It provides an intuitive drag-and-drop interface, making it accessible to both technical and non-technical users. Tableau enables users to gain insights from data quickly and efficiently, facilitating data-driven decision-making across organizations.

**Connecting to Data Sources:**

Tableau offers seamless integration with various data sources, including databases, Excel files, and cloud-based services. Users can connect to data sources using Tableau's built-in connectors or by establishing direct connections using custom SQL queries. Tableau supports live connections.

**Creating Visualizations:**

In Tableau, users can create a wide range of visualizations to explore and analyse data effectively. Common visualizations include:

**Bar Charts**: Used to compare categorical data.

**Line Graphs**: Used to visualize trends over time.

**Scatter Plots**: Used to identify relationships between two numerical variables.

**Maps**: Used to display geographical data and spatial relationships.

Tableau provides a drag-and-drop interface for building visualizations, allowing users to customize the appearance, size, and formatting of charts and graphs easily.

**Interactive Dashboards:**

Tableau enables users to combine multiple visualizations into interactive dashboards, allowing users to explore data dynamically. Users can add filters, parameters, and actions to dashboards to provide interactivity and drill-down capabilities.

**Sharing and Collaboration:**

Tableau provides several options for sharing visualizations with others and collaborating on projects. Users can publish their Tableau workbooks and dashboards to Tableau Server or Tableau Online, where they can be accessed by authorized users via web browsers or Tableau Desktop.

**Policies and Procedure-Task (1)**

**Data Privacy Policy:**

**Definition**: Data privacy policies govern the collection, use, and sharing of personal and sensitive information to protect individuals' privacy rights.

Adhering to data privacy policies ensures that personal information within "The Wealth of Nations" dataset is handled ethically and legally, minimizing the risk of unauthorized access, misuse, or breaches.

**Data Security Policy:**

**Definition:** Data security policies outline measures and procedures to safeguard data against unauthorized access, disclosure, alteration, or destruction.

Implementing data security measures is essential to protect the confidentiality, integrity, and availability of "The Wealth of Nations" dataset, reducing the risk of data breaches, cyberattacks, and data loss.

**Data Governance Policy:**

**Definition**: Data governance policies establish guidelines for managing and controlling data assets, ensuring data quality, consistency, and compliance with regulations.

Adhering to data governance policies promotes transparency, accountability, and trustworthiness in data management practices, facilitating accurate analysis and decision-making based on "The Wealth of Nations" data.

**Ethical Data Use Policy:**

**Definition:** Ethical data use policies define acceptable and responsible behaviour when working with data, including respect for individuals' rights, diversity, and cultural sensitivity.

Following ethical data use principles promotes integrity, fairness, and social responsibility in data analysis, fostering trust among stakeholders and mitigating ethical risks associated with "The Wealth of Nations" data.

**Regulatory Compliance Policy:**

**Definition:** Regulatory compliance policies ensure adherence to relevant laws, regulations, and industry standards governing data management, privacy, and security.

Complying with regulatory requirements minimizes legal and financial risks associated with non-compliance, such as penalties, fines, and reputational damage, while also ensuring the lawful use of "The Wealth of Nations" data.

**As a data analyst, awareness of these policies is crucial for several reasons:**

**Legal Compliance:** Ensuring compliance with data protection laws and regulations reduces the risk of legal liabilities and sanctions.

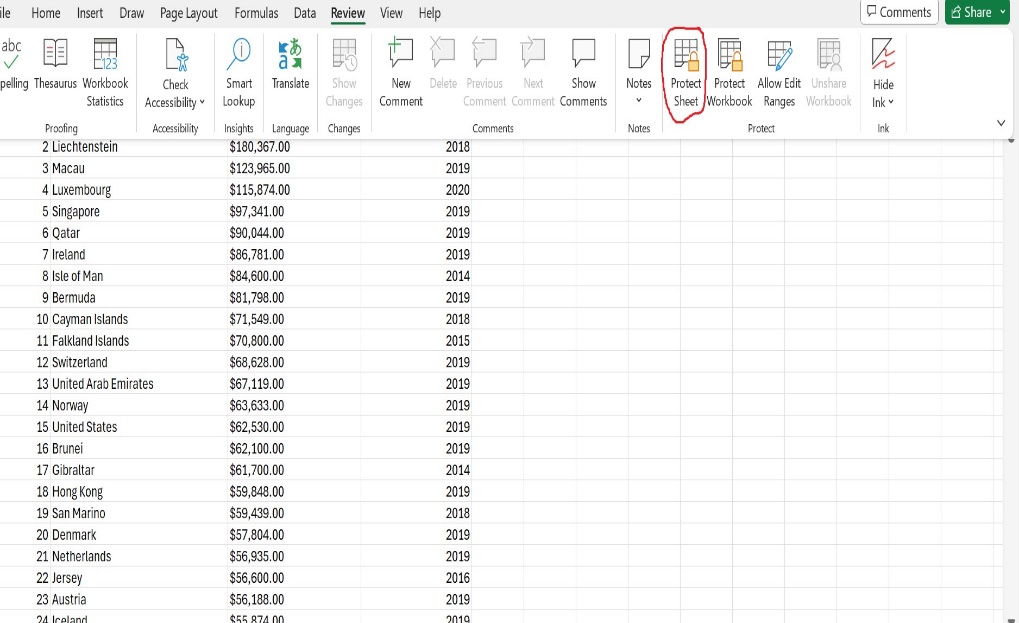
**Data Integrity:** Adhering to data governance and security policies maintains the integrity and reliability of analysis results and conclusions.

**Trust and Reputation:** Upholding ethical and professional standards builds trust with stakeholders and enhances the reputation of both the analyst and the organization.

**Risk Mitigation:** Following established policies helps mitigate risks associated with data breaches, privacy violations, and ethical misconduct.

**GDP Sheet** - TASK (2)

1. Setting Password to protect workbook



A screenshot of a spreadsheet

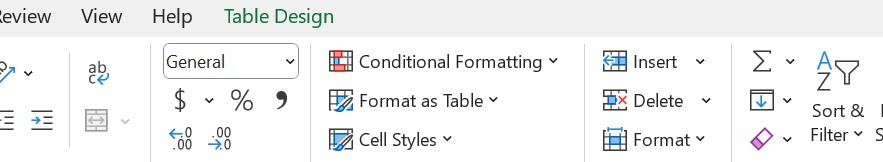
Description automatically generated

A screenshot of a computer

Description automatically generated

Here we have protected the work book by going to review page on Ribbon tab go to protect heading and click on protect workbook and enter the password click ok and it will open the re-enter the password dialogue box re-enters the password and click ok.

1. Changing the format and turning the sheet to table and displaying 2019 information.



A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

We need to change the column C from $ to £ from home tab go to number field change the Currency to £.

Change the same sheet to table format go to insert tab click on tables and click ok then filter the year to 2019 by clicking the arrow down button on column D(year of information) de select all and select 2019.

1. Create Edit the chart to display Rand, Country, GDP-per Capital (PPP)

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

After Editing the GDP Table select the Rank, Country, GDP-per capital(ppp) columns go to insert tab and click on charts choose bar chart or suitable chart now we need to add or edit the Axis labels once the chart is created a new table will appear Chart Design go to add chart elements and edit the X,Y axis titles and chart titles.

1. Moving the chart to new sheet.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Now the created chart need to move to new sheet right click on the chart and click on move chart a pop up will appear to choose where we want the chart to move click on new sheet and click ok.

1. Creating a new bar chart by sorting the GDP sheet to top 20 highest country.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Here we have sort the GDP sheet to top 20 highest country using sort function under number filter choose top 10 option and change to 20 then top 20 country will be appear.

Select the sorted data and choose the bar chart for the chart section in insert tab and place the chart under the data using the colour option in the home tab under font we can change the background colour.

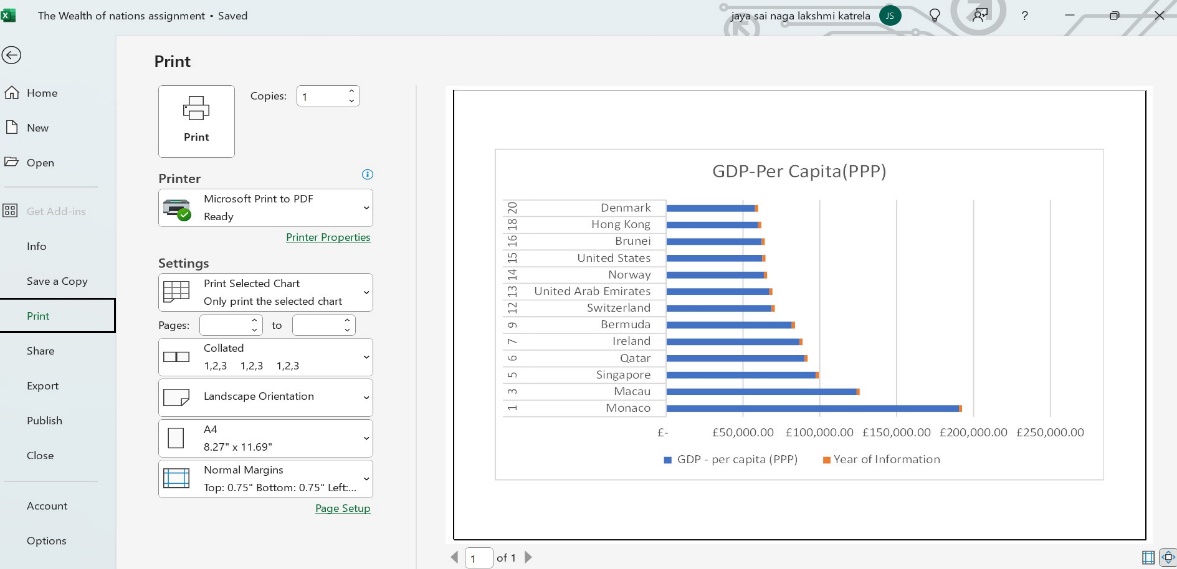
**Creating the Macros**

1. Creating print, copy, save macro buttons.

A screenshot of a computer

Description automatically generated

To create the Macro first we need to start record the macro under Developer tab in code section click on record macro the pop up will appear for the record macro and enter the macro name and we can enter the short cut key and which work book you want to store the macro choose the workbook we can enter the description to describe what the macro is doing then click ok.



A screenshot of a computer

Description automatically generated

Perform the print process manually and stop the recording now the Macro was created for the print option now create the button and assign the task.

A screenshot of a computer

Description automatically generated

To assign the macro first create button from developer tab under control section go to insert and click on Form control and choose first shape and draw the shape.

A screenshot of a computer

Description automatically generated

Edit the text by right clicking the button and choose edit text option.

A screenshot of a computer

Description automatically generated

Now assign the macro by right clicking the button choose Assign Macro and click on the print option from the list and choose the workbook where we have recorded the macro and select ok.

Same process for copy and save first need to record the macro and do the process manually and stop the recording and create the button and assign the macro for the buttons.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

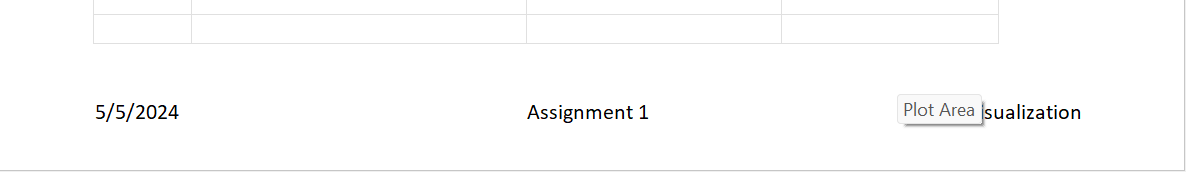
Once the macro is created test them and change the copied macro sheet with GDP(Gross Domestic Product).

Now add the header and footer to the table.

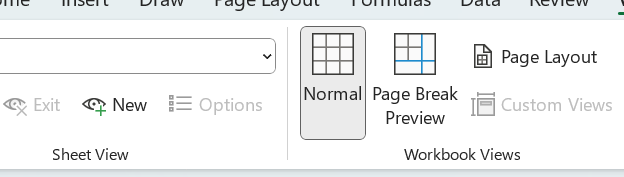
A screenshot of a computer

Description automatically generated

To add header and footer go to view tab under workbook view select page layout and edit the header and footer section.



Once the header and footer added click on normal view.



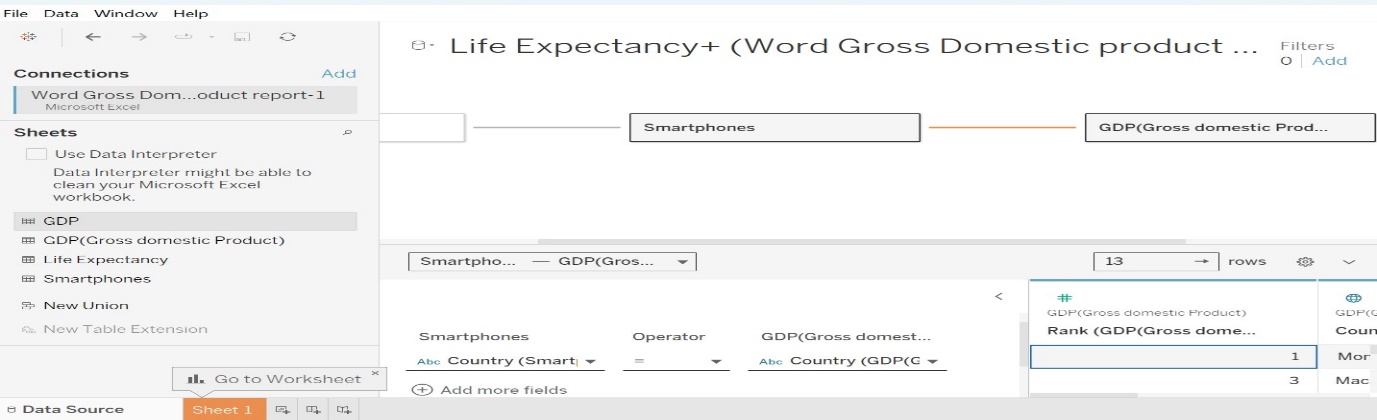
**Creating Charts in Tableau**

Import the excel sheet with we have created in to tableau public and set the relation between the sheets which is common and check the data types now we can start creating the charts.

A screenshot of a computer

Description automatically generated

Open file and create relationship



Click on sheet 1 and create the charts

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Once the chat is created publish the file navigate to file tab click on save to tableau public and click ok it will open in your tableau profile.