

SUMMARY: DATA ANALYSIS-II

(Univariate Analysis)

SESSION OVERVIEW:

By the end of this session, you will be able to:

- Understand different aspects of charts and pivot charts in Excel
- Understand how the dashboards are created using pivot tables and charts.
- Understand the introduction of univariate analysis.

KEY TOPICS AND EXAMPLES:

1. Understanding charts in Excel: (30 min)

a. What are charts?

Charts in Excel are visual representations of data that allow users to interpret and analyze information more easily. They display data points, trends, and patterns in a graphical format, making it simpler to understand complex datasets. Excel offers various types of charts, such as column charts, line charts, pie charts, scatter plots, and more, each serving different purposes and suited to different data types.

b. Importance of charts in Excel:

Visualizing Data: Charts provide a visual representation of numerical data, making it easier to interpret and understand trends, patterns, and relationships.

Comparison: Charts allow users to compare data points, categories, or trends visually, facilitating quick analysis and decision-making.

Trend Identification: By plotting data points over time or other variables, charts help identify trends, cycles, and patterns that may not be apparent in raw data.

Presentation: Charts enhance the presentation of data in reports, presentations, and dashboards, making them more engaging and understandable for stakeholders.

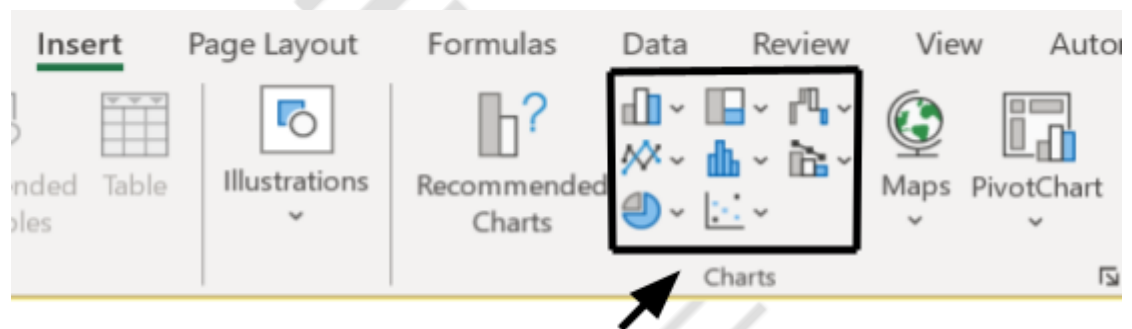
Data Analysis: Charts aid in data analysis by providing insights into distributions, outliers, correlations, and other statistical properties of the data.

Communication: Charts help communicate key findings and insights to others effectively, enabling better collaboration and communication among team members and stakeholders.

c. Different types of charts in Excel:

([Reference Dataset](#))

- **Column Chart:** Represents data using vertical bars, ideal for comparing values across categories.
- **Bar Chart:** Similar to column charts but with horizontal bars, useful for comparing values across categories.
- **Line Chart:** Displays data as a series of data points connected by straight lines, suitable for showing trends over time.
- **Pie Chart:** Represents data as a circle divided into slices, with each slice representing a proportion of the whole.
- **Area Chart:** Similar to line charts but with the area below the line filled in, useful for comparing total values over time.
- **Scatter Plot:** Displays individual data points as dots on a graph, helpful for showing relationships between two variables.
- **Bubble Chart:** Similar to scatter plots but with bubbles of different sizes representing the third variable.
- **Histogram:** Represents the distribution of numerical data in bins or intervals, useful for understanding data distributions.
- **Radar Chart:** Displays data in a circular manner with multiple axes, helpful for comparing multiple variables across categories.
- **Combo Chart:** Combines two or more chart types into a single chart, allowing for better visualization of complex data.



- We can find different types of charts in the “Insert” section in the Excel ribbon.
 - Once we click insert, there will be an option with charts along with some icons of charts as visible in the above image.
- **Map charts:**
 - **Geographical Data Analysis:** Map charts allow users to analyze geographical data, such as sales data by region, population density, or distribution of resources.
 - **Regional Performance Comparison:** Businesses can use map charts to compare performance metrics, such as sales revenue or market share, across different regions.

- **Educational Purposes:** Map charts are valuable tools for educational purposes, allowing students to learn about geography, demographics, and social trends.
- **Public Health and Epidemiology:** Health organizations and researchers can use map charts to track and visualize the spread of diseases, outbreaks, and epidemics.

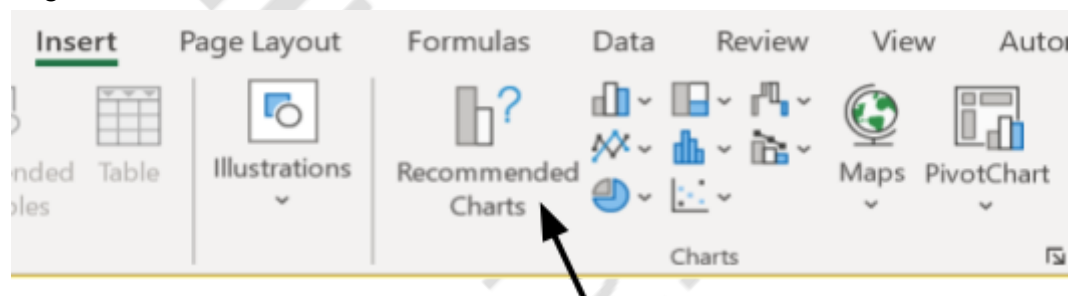
c. Steps involved in creating charts:

- **Select Data:** Highlight the data range in your Excel worksheet that you want to include in the chart, including the column and row labels if applicable.
- **Insert Chart:** Navigate to the "Insert" tab on the Excel ribbon.
- **Choose Chart Type:** Click on the "Charts" group and select the desired chart type from the options available, such as column, line, pie, bar, or scatter plot.
- **Customize Chart:** Once the chart is inserted, customize it according to your preferences.

d. Recommended charts:

([Reference dataset](#))

Recommended charts is an AI feature in MS Excel that helps us with the ease of visualization by recommending charts with respect to the selected data range.



Steps:

- In the Insert section, we can find recommended charts.
- Once we click on the recommended chart option, this dialogue box appears suggesting different types of charts.
- Once we select the desired chart, a chart gets created corresponding to the dataset or the pivot table.

e. Pivot charts:

Uses:

- **Visualizing Pivot Table Data:** Pivot charts provide a visual representation of the data summarized in a pivot table. They allow users to quickly grasp trends, patterns, and insights that may not be as apparent when viewing the raw data.
- **Dynamic Filtering and Analysis:** Pivot charts are interactive, allowing users to filter and analyze data dynamically. Users can change filters, slicers, and other settings to explore different aspects of the data and drill down into specific details.

- **Comparative Analysis:** Pivot charts facilitate comparative analysis by enabling users to compare different categories, series, or groups within the data. Users can easily identify trends, disparities, and relationships between various data points.
- **Presenting Insights and Reports:** Pivot charts are effective tools for presenting insights and reports to stakeholders, managers, or clients. They provide a visually appealing and concise way to communicate key findings, trends, and analysis results.

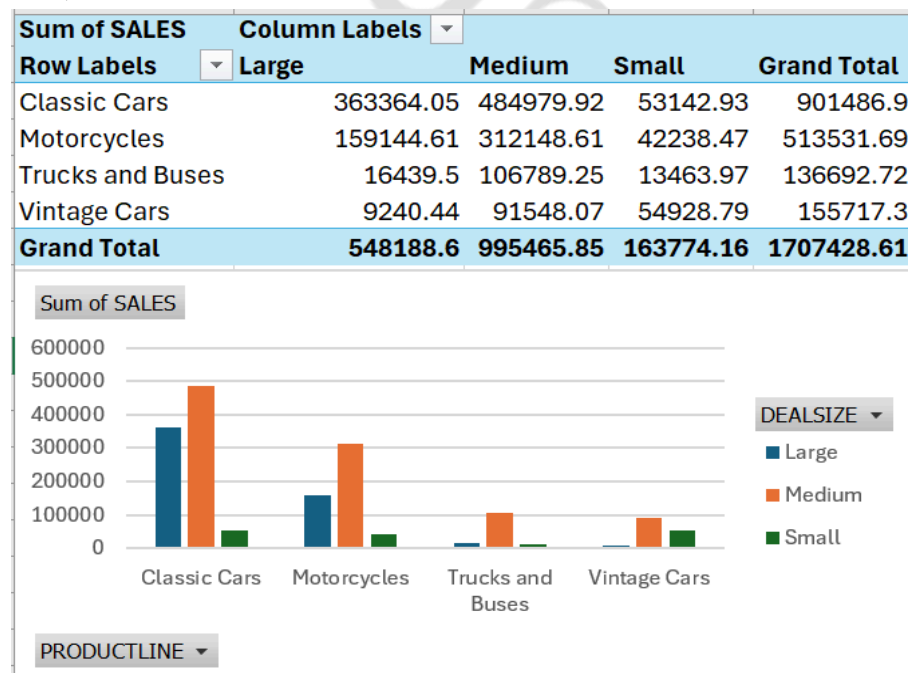
Steps:

- Ensure your data is organized properly with clear headers and no empty rows or columns.
- Select your data range, go to the "Insert" tab in Excel, and click on "PivotTable". Choose where you want the pivot table to be placed (either a new worksheet or an existing one) and click "OK".
- In the PivotTable Field List pane, drag the fields you want to analyze into the "Rows", "Columns", and "Values" areas.
- With your pivot table selected, go to the "Insert" tab and click on "PivotChart". Choose the chart type you want to create and click "OK".
- Once the pivot chart is created, you can customize it by adding titles, axis labels, legends, and formatting options.

Sample example:

[\(Reference Dataset\)](#)

Once you follow the above steps, you will be successful in creating charts of your desired categories as shown below. This is just a sample chart (Column chart).



So here we can easily conclude from the chart that the **highest sales are from medium-sized deals of classic cars.**

Similarly, insights can be derived from a dataset using different types of charts in Excel.

FUN TIME: (30 mins) (Dataset)

With the help of the above topics, we will be creating a small dashboard using this [dataset](#).

The dataset represents the purchases of bikes and consists of customer demographic details like income, gender, age, commutation distance, marital status, and many more.

Here with the help of pivot tables, slicers, and charts, we will be understanding the behavior of the customers who are more inclined towards purchasing bikes.

This activity will be conducted by the instructor during the class.

2. Understanding different types of analysis: (10 mins)

a. Different types of data:

i. Numerical Data:

1. **Discrete Data:** This type of data consists of countable, whole numbers. Examples include the number of children in a family, the number of defective items in a batch, or the number of customers visiting a store.
2. **Continuous Data:** This type of data can take any numeric value within a specific range, including decimal values. Examples include height, weight, temperature, or time measurements.

ii. Categorical Data:

1. **Nominal Data:** This type of data consists of categories or groups that have no inherent order or ranking. Examples include gender (male, female), eye color (blue, green, brown), or type of car (sedan, SUV, hatchback).
2. **Ordinal Data:** This type of data consists of categories or groups that have a natural order or ranking. Examples include educational levels (high school, bachelor's, master's, doctorate), socioeconomic status (low, middle, high), or customer satisfaction ratings (poor, fair, good, excellent).

b. Different types of data analysis:

i. Univariate analysis:

Univariate analysis involves the examination of a single variable at a time. It is used to describe and summarize the characteristics of a single variable, such as its central tendency (mean, median, mode), dispersion (range, variance, standard deviation), and distribution (normal, skewed, etc.).

Univariate analysis techniques include frequency distributions, histograms, box plots, and summary statistics.

ii. **Bivariate analysis:**

Bivariate analysis involves the examination of the relationship between two variables. It helps to understand how one variable is associated with or influenced by another variable. Bivariate analysis techniques include scatter plots, correlation analysis (e.g., Pearson's correlation coefficient), and contingency tables.

iii. **Multivariate analysis:**

Multivariate analysis involves the examination of the relationships among three or more variables simultaneously. It allows for a more comprehensive understanding of complex relationships and the ability to control for confounding variables.

3. Brief understanding of Univariate analysis:

a. **What is univariate analysis:**

Univariate analysis is a statistical technique used to analyze and describe a single variable in a dataset. It involves summarizing and exploring the characteristics of one variable at a time, without considering any relationships with other variables. The goal of univariate analysis is to understand the distribution, central tendency, and dispersion of a single variable.

b. **Different types of univariate analysis:**

i. **Continuous univariate analysis:**

Continuous univariate analysis in Excel involves examining the distribution and characteristics of a single continuous variable within a dataset. This type of analysis is commonly used to explore the characteristics of numerical data, such as heights, weights, or temperatures.

To perform continuous univariate analysis, the data is first prepared by organizing it in a single column. Summary statistics, such as mean, median, standard deviation, minimum, and maximum, are calculated to describe the central tendency and spread of the data.

visualizations such as histograms and box plots are created to visualize the distribution of the data and identify any patterns or outliers.

Histograms display the frequency distribution of data in different intervals, while box plots summarize the distribution by showing quartiles, median, and any potential outliers.

ii. **Categorical univariate analysis:**

Categorical univariate analysis involves the examination of a single categorical variable within a dataset. Unlike continuous variables, which represent numerical measurements, categorical variables represent distinct

categories or groups. The goal of categorical univariate analysis is to understand the distribution of categories within the variable and assess their frequencies or proportions.



codingninjas