



MASTER OF COMPUTER APPLICATIONS

SEMESTER 1

DATA VISUALIZATION

Unit 2

Excel in Data Visualization

Table of Contents

SL No	Topic	Fig No / Table / Graph	SAQ / Activity	Page No
1	Introduction	-	-	3
1.1	Learning Objectives	-	-	
2	Introduction to Excel	-	1	4-9
2.1	Overview of Microsoft Excel	-	-	
2.2	Spreadsheet Basics	-	-	
3	Data Types	-	2	10-16
3.1	Understanding Data Types in Excel	-	-	
3.2	Formatting Data in Excel	-	-	
4	Visualization Capabilities	-	3	17-23
4.1	Creating Charts and Graphs in Excel	-	-	
4.2	Formatting Charts in Excel	-	-	
5	Summary	-	-	24
6	Glossary	-	-	25-28
7	Terminal Questions	-	-	28-29
8	Answers	-	-	29-30
9	Suggested Books and E-References	-	-	31

1. INTRODUCTION

In this unit, we focus on mastering Excel for data visualisation. If you've ever been in a position where you need to make sense of large sets of data or present data-driven findings, you know that a well-crafted chart or graph can make all the difference. Excel is not merely a tool for organising your data; it's essential for turning raw data into actionable insights through visual representation.

Understanding Excel's capabilities in data visualisation is vital for various reasons. It simplifies complex data sets and reveals essential patterns, trends, and relationships. This aids in data analysis and decision-making processes, making you an asset in any professional setting. Whether in a management role, a researcher, an analyst, or just starting your career, knowing how to visualise data effectively will give you a competitive edge.

Excel's user-friendly interface lets you quickly grasp how to create impactful charts and graphs, enabling you to present complex data in an easily understandable format. This skill is invaluable in today's data-driven world, where making timely and informed decisions can significantly impact the success of projects and organisations.

In this unit, the students will learn everything from the basics of Excel, like understanding different data types and formatting options, to advanced topics, like creating various charts and customising them for maximum impact. These skills are not just theoretical but efficient and applicable in your professional life.

1.1 Learning Objectives

By the end of this unit, the students will be equipped with the skills to:

- ❖ *Evaluate the effectiveness of various Excel chart types in representing different data sets to make informed decisions.*
- ❖ *Apply basic and advanced formatting techniques to improve data visualisation in Excel spreadsheets.*
- ❖ *Analyse raw data to determine the most appropriate visualisation method in Excel.*
- ❖ *Recall critical features and tools in Excel that are essential for data visualisation.*

2. INTRODUCTION TO EXCEL

Microsoft Excel is a versatile, powerful, and widely used spreadsheet application that forms a cornerstone of Microsoft's Office Suite. It was first introduced in 1985 and has since undergone significant changes, catering to many desktop and online platforms, including Windows, macOS, Android, and iOS. Excel is often considered the gold standard in spreadsheet software, not merely for its abilities to organise and calculate data but for its comprehensive range of functions that extend to data visualisation, statistical analysis, and even rudimentary forms of machine learning.

2.1 Overview of Microsoft Excel

At its core, Microsoft Excel is a spreadsheet application that allows users to store, manipulate, and analyse data in a tabular format. A standard Excel workbook consists of one or more 'worksheets,' each containing rows and columns that intersect to form 'cells.' *Cells are the primary unit for data entry and can hold different data types like numbers, text, formulas, and references to other cells.*

Excel's ubiquity arises from its dual nature. It offers straightforward data storage and operations for the beginner or casual user. For the advanced user, it provides a depth of functionality ranging from pivot tables to VBA (Visual Basic for Applications) scripting, providing a sophisticated data manipulation and analysis platform.

Importance in the Modern World

Given the sheer volume of data generated today, the importance of a tool that can efficiently manage and make sense of this data cannot be overstated. Excel comes into play in various sectors:

- **Business:** For financial forecasting, inventory management, and market analysis.
- **Healthcare:** In statistical analysis, patient data management, and reporting.
- **Academia:** For research, data compilation, and even administrative tasks.
- **Individual Use:** For budgeting, planning events, and tracking personal goals.

Features of Excel

Excel comes loaded with features that make it a versatile tool for multiple tasks:

1. **Data Storage and Management:** It can handle large datasets, offering various ways to sort, filter, and search data.
2. **Mathematical Functions:** Over 400 functions like SUM, AVERAGE, and VLOOKUP for different mathematical operations.
3. **Data Visualisation Tools:** Options to create charts, graphs, and even interactive dashboards.
4. **Pivot Tables:** For summarising, analysing, and presenting your data.
5. **Macro and Scripting:** With VBA, you can automate tasks and perform complex operations.
6. **Collaboration Features:** Cloud-based versions allow for real-time collaboration.

User-Friendly Interface

One of Excel's most vital points is its user-friendly interface. Most features can be accessed from the ribbon at the top of the screen, divided into tabs like 'Home', 'Insert', 'Formulas', etc., each holding a set of related tools. Additionally, Excel offers various templates for budgets, calendars, invoices, and other standard documents, making it easier for new users to get started.

Limitations

However, it's important to note that Excel is not without its limitations:

- **Size Restrictions:** There's a maximum limit of 1,048,576 rows and 16,384 columns per worksheet.
- **Not a Database:** While it can perform some functions, it lacks the relational capabilities of a proper database like SQL.
- **Learning Curve:** Some advanced features have a steep learning curve and may require specialised training.

Versions and Pricing

Excel has evolved to meet user demands, offering various versions like Excel 2019, Excel 2016, and Excel for Microsoft 365, with each subsequent version featuring incremental improvements and features. Pricing varies depending on whether you purchase Excel as

standalone software or as part of the Microsoft Office Suite and whether it's for individual or business use.

2.2 Spreadsheet Basics

A spreadsheet is a digital canvas for organising, analysing, and storing data in a tabular format. While the idea of a spreadsheet predates digital technology, having existed in paper form for accounting and data management, the digital versions—like Microsoft Excel—have transformed how businesses, academics, and individuals work with data.

What is a Spreadsheet?

Essentially, a spreadsheet comprises rows and columns that intersect to form cells. Each cell serves as a data point, a number, text, date, formula, or even a hyperlink to another resource. The beauty of spreadsheets is their flexibility; they can serve as a simple calculator or a complex data analysis tool, depending on the user's needs.

Core Elements of a Spreadsheet

- **Cell:** The basic building block of a spreadsheet, identified by the intersection of a row and a column.
- **Row:** A horizontal sequence of cells, typically identified by a number (e.g., Row 1, Row 2, etc.).
- **Column:** A vertical sequence of cells, usually identified by a letter or sequence of letters (e.g., Column A, Column B, etc.).
- **Worksheet or Sheet:** One page of a spreadsheet file, which can hold data, charts, and other elements.
- **Workbook:** The actual Excel file that contains one or more worksheets.

Data Entry Basics

Entering data into a spreadsheet is as straightforward as clicking on a cell and typing. Here's how to manage data entry:

1. *Simple Entry:* Click on a cell and start typing to enter data.
2. *Bulk Entry:* For entering multiple rows or columns of data, you can copy and paste or drag to fill cells.

3. *AutoFill Feature*: Excel can automatically populate cells based on the data pattern it detects.

Formatting Cells

Formatting is a powerful feature in Excel that enables users to change the appearance of cells without altering the underlying data. Here's how to format cells effectively:

1. **Numeric Formatting**: Changing how numbers are displayed, like currency, percentage, or scientific notation.
2. **Text Formatting**: Altering the font, style, and alignment of text.
3. **Cell Style**: Applying predefined styles, including borders and background colours.

Navigating a Spreadsheet

1. **Scroll Bars**: Use the horizontal and vertical scroll bars to navigate.
2. **Arrow Keys**: Use arrow keys to move cell by cell.
3. **Page Up/Down**: Jump an entire screen up or down.
4. **Ctrl + Arrow Key**: Quickly move to the edge of the data region.

Functions and Formulas

Formulas and functions are among the most powerful tools in Excel, allowing for various calculations:

Basic Arithmetic: Formulas can be as simple as =A1+B1 for addition.

Built-In Functions: Excel comes with hundreds of predefined functions like SUM(), AVERAGE(), and VLOOKUP().

Data Sorting and Filtering

These features help in managing large sets of data:

1. **Sort**: One can sort data in ascending or descending order.
2. **Filter**: This lets one to view specific rows based on criteria without altering the table's structure.

Data Analysis Tools

Excel offers numerous data analysis tools:

1. **Pivot Tables:** For advanced data analysis and summarisation.
2. **Data Validation:** For setting rules determining what can be entered into a cell.
3. **What-If Analysis:** Tools like Goal Seek and Scenario Manager for predictive analysis.

Protecting Your Spreadsheet

1. **Cell Locking:** To prevent unauthorised changes.
2. **Password Protection:** For securing the entire workbook or individual sheets.
3. **Read-Only:** To allow viewing but not editing.

Cloud-Based Sharing and Collaboration

Modern versions of Excel allow cloud-based collaboration, a critical feature for team projects. One can share the entire workbook or individual sheets and work on them in real time with team members.

Importing and Exporting Data

Excel provides features to import data from various formats like .csv, .txt, or a database. Conversely, you can export Excel data to multiple formats for other applications.

Key Shortcuts and Tips

Ctrl + C and Ctrl + V for Copy and Paste.

Ctrl + Z to Undo an action.

Ctrl + Shift + Arrow Key to select an entire range of cells. Further use.

SELF-ASSESSMENT QUESTIONS – 1

1. What does a cell in an Excel spreadsheet represent?
 - a) A sheet
 - b) An intersection between a row and a column
 - c) A chart
 - d) A file
2. What is the first column label in Excel?
 - a) 0
 - b) A
 - c) 1
 - d) AA
3. How is a range of cells usually represented?
 - a) A1C3
 - b) A1:C3
 - c) A1_C3
 - d) A1-C3
4. Which of the following is NOT a way to navigate through cells in Excel?
 - a) Arrow keys
 - b) Mouse clicks
 - c) Swiping on the touchpad
 - d) Scrolling the mouse wheel
5. What does 'CTRL + Space' do in Excel?
 - a) Deletes a cell.
 - b) Select the entire column.
 - c) Moves to the next sheet.
 - d) Saves the document.

3. DATA TYPES

Data types are the building blocks of any spreadsheet in Excel, determining the kind of information that can be stored in cells. Understanding these data types is crucial for effectively organising, analysing, and visualising your data. This section provides an in-depth look at the various data types in Excel, how they function, and how they impact your work.

3.1 Understanding Data Types in Excel

Data types are vital for various reasons:

1. **Efficient Storage:** Knowing the correct data type minimises memory usage.
2. **Data Integrity:** Using the correct data type ensures the accuracy of calculations and analyses.
3. **Usability:** Specific data types can dictate what operations or functions can be performed on them.

Core Data Types in Excel

Excel supports several core data types:

1. **Text (String):** Alphanumeric characters, including letters, numbers, and special characters, but usually not used for calculations.
2. **Number:** Numeric values that can be used in arithmetic operations.
3. **Date and Time:** Encodes dates and times in a specific format.
4. **Boolean:** Represents True or False conditions.
5. **Error Types:** Includes types like #DIV/0!, #N/A, #NAME?, etc., indicating errors in formulas.
6. **Particular Types** include hyperlinks, rich text, and other non-standard data types.

Text (String) Data Type

Text or String is a commonly used data type in Excel. It is helpful for:

1. **Naming Columns:** E.g., "Product Name," "Employee ID," etc.
2. **Data Labelling:** E.g., "Yes," "No," "Pass," "Fail," etc.
3. **Annotations:** To add comments or descriptions.

Limitations

- It cannot be directly used in most arithmetic calculations.
- Sorting text is alphabetical rather than numerical.

Number Data Type

Numbers are the workhorse of Excel, involved in almost every calculation.

1. **Integers:** Whole numbers without decimal places.
2. **Decimal Numbers:** Numbers with decimal places.
3. **Scientific Notation:** Useful for very large or small numbers.

Types of Numeric Operations

1. **Arithmetic Operations:** Such as addition, subtraction, etc.
2. **Statistical Functions:** Such as AVERAGE, SUM, MAX, etc.

Limitations

- Rounding errors in floating-point calculations.
- It is limited by the maximum and minimum values Excel can handle.

Date and Time Data Type

Excel has specific formats for Date and Time:

1. **Date:** In the format of "MM-DD-YYYY," "DD-MM-YYYY," etc.
2. **Time:** Represented in "HH:MM: SS AM/PM" format.

Features

1. **Date and Time Functions:** Such as TODAY(), NOW(), DAY(), etc.
2. **Date Arithmetic:** This is calculating the number of days between two dates.

Limitations

- Susceptible to locale and format settings.
- Leap years and daylight saving time can cause complications.

Boolean Data Type

Boolean data types represent TRUE or FALSE conditions and are used in logical operations like AND, OR, and NOT.

Error Types

These represent errors in your spreadsheet and often stem from incorrect formulas or data types. Recognising and understanding these is critical for debugging.

1. **#DIV/0!:** Division by zero.
2. **#N/A:** Value not available.
3. **#NAME?:** Excel doesn't recognise a name or function.

Special Types

These include:

1. **Hyperlinks:** Clickable links stored in a cell.
2. **Rich Text:** Text with different kinds of formatting within the same cell.

Advanced-Data Types in Excel 365

With Excel 365, Microsoft introduced new data types powered by its knowledge graph.

These include:

1. **Stocks:** Real-time stock data.
2. **Geography:** Geographical data like country names, cities, etc.
3. **Custom Data Types:** Created via Power Query or through formulas.

These advanced data types allow users to perform unique analyses that were not possible in previous versions of Excel.

3.2 Formatting Data in Excel

Excel is not just about performing calculations and creating graphs; the presentation of your data is equally important. Formatting your Excel sheets improves readability, comprehension, and usability. This section delves into the vast array of options that Excel provides for data formatting, clarifying how and when to use them.

Importance of Formatting

The importance of data formatting cannot be overstated for several reasons:

1. **Readability:** Well-formatted data is easier to read and interpret.
2. **Data Integrity:** Proper formatting can prevent errors by making issues more visible.
3. **Professionalism:** Well-formatted sheets are perceived as more professional and trustworthy.

4. **Efficiency:** Formats can be applied conditionally or in bulk, saving time in data analysis.

Basic Formatting Options

Excel offers a variety of basic formatting options, such as:

1. **Font Styling:** Includes font type, size, and colour.
2. **Cell Fill:** Background color for the cell.
3. **Borders:** Line styles that separate cells from each other.
4. **Text Alignment:** Horizontal or vertical alignment of text within cells.
5. **Number Formatting:** Such as percentages, currency, date and time.

Font Styling

Font styles should be consistent throughout the spreadsheet to ensure readability. Options include:

- **Bold:** For headings or to emphasise data.
- **Italics:** For sub-headings or annotations.
- **Underline:** To mark essential data points.

Best Practices

- Use sans-serif fonts for digital documents.
- Use a font size of 11 or 12 for body text.

Cell Fill and Borders

Cell background colours can be used to group or highlight data. Borders help in visually separating data.

Best Practices

- Use subtle colours for cell fill.
- Use strong borders for tables and lighter borders within.

Text Alignment

Alignment can be both horizontal and vertical.

- **Left Align:** For text data.
- **Right Align:** For numerical data.
- **Center Align:** For headings.

Best Practices

- Consistency is key.
- Use indentation for sub-categories.

Number Formatting

Excel offers a myriad of options to format numbers. Some popular options include:

- **General:** No specific formatting.
- **Number:** For general display of numbers.
- **Currency:** For monetary values.
- **Percentage:** Multiplies the cell value by 100 and displays it with a % symbol.

Advanced Formatting Options

Advanced formatting can be more dynamic and interactive:

1. **Conditional Formatting:** Changes the formatting based on certain conditions.
2. **Data Bars or Color Scales:** Adds a background colour representing data magnitude.
3. **Custom Formatting:** Created through the Format Cells dialog box.
4. **Themes:** Apply a unified colour, font, and effects scheme to the workbook.

Conditional Formatting

Excel allows you to define conditions under which specific formatting options will be applied. This is especially useful for dashboarding and data visualisation.

Examples

- Highlight cells that are above a specific value.
- Colour-code percentages to easily spot high or low values.

Best Practices

1. Limit the number of conditions to avoid confusion.
2. Use contrasting colours for better visibility.

Data Bars and Color Scales

Data Bars and Color Scales are a type of conditional formatting that visually represent data.

Best Practices

- Use data bars for representing ranges.
- Use colour scales for heat maps.

Custom Formatting

Excel allows custom number formatting through the "Format Cells" dialog. You can even use it to create particular data types or conditions that are not natively supported.

Themes

Themes are excellent for maintaining consistency throughout a giant workbook, especially if you're working collaboratively.

Formatting for Data Types

Different data types often require specific formatting. For instance, Date and Time should be consistently formatted, and Numbers can have decimal places standardised for easier reading and calculating.

SELF-ASSESSMENT QUESTIONS – 2

6. What does 'Conditional Formatting' do in Excel?
 - a) It hides unnecessary data.
 - b) It formats cells based on specific conditions.
 - c) It creates a new sheet.
 - d) It adds new rows automatically.
7. What function can you use to join two text cells in Excel?
 - a) ADD
 - b) CONCATENATE
 - c) SUM
 - d) MULTIPLY
8. What is the purpose of 'Wrap Text' in Excel?
 - a) It enlarges the text size.
 - b) It wraps text within a cell.
 - c) It changes the text to italics.
 - d) It deletes extra spaces.
9. What does the 'Merge Cells' option do?
 - a) It splits a single cell into multiple cells.
 - b) It adds a border around cells.
 - c) It combines multiple cells into a single cell.
 - d) It duplicates the selected cells.
10. How do you set the background colour of a cell?
 - a) Using the 'Background' tab
 - b) Using the 'Font' dialog box
 - c) Using the 'Fill Color' option
 - d) Using the 'Cell Size' option

4. VISUALISATION CAPABILITIES

Visual representation of data often communicates insights more effectively than tables or text. Excel has robust capabilities to create a variety of charts and graphs.

Why Visualise Data?

Data visualisation has several benefits:

1. **Quick Insights:** It enables faster interpretation of data.
2. **Complex Data Simplification:** Makes understanding complicated data structures easier.
3. **Enhanced Communication:** Aids in presenting findings in a more digestible way.
4. **Decision-making:** Enables quicker and better-informed decisions.

4.1 Creating Charts and Graphs in Excel

Creating a chart in Excel usually involves:

1. **Selecting Data:** Choose the data that you want to represent visually.
2. **Choosing the Right Chart Type:** Deciding which chart represents your data best.
3. **Inserting the Chart:** Using Excel's Insert Tab to add the chart.
4. **Customisation:** Tweaking the chart elements to suit your needs.

Types of Charts and Graphs

Excel offers a myriad of chart types, each with its unique use cases:

1. **Column Chart:** Used for showing data changes over some time or for illustrating comparisons among items.
2. **Line Chart:** Ideal for displaying trends over periods.
3. **Pie Chart:** Useful for showing a part-to-whole relationship.
4. **Bar Chart:** Suitable for comparing quantities of different categories.
5. **Scatter Plot:** Great for illustrating the correlation between two variables.
6. **Area Chart:** Like a line chart, the area under the line is filled in to indicate volume.

Column Chart

A Column chart is one of the most frequently used chart types. They are excellent for showing variations in the value of an item over time.

Best Practices

- Limit the number of columns to avoid clutter.
- Use a horizontal axis for time periods and a vertical axis for values.

Line Chart

Line Charts are crucial for showing trends over time.

Best Practices

- Use markers for critical data points.
- Limit the number of lines for clarity.

Pie Chart

Pie Charts are excellent for representing proportional or percentage data.

Best Practices

- Limit the number of slices to 5-7 for readability.
- Use distinct colours for each slice.

Bar Chart

Bar Charts are essentially Column Charts turned on their side. They are helpful when you have long category names that need more display room.

Best Practices

- Consider using a horizontal bar chart if the category names are long or if you have more than seven categories.

Scatter Plot

Scatter Plots are perfect for showing correlations between two sets of values.

Best Practices

- Use a trendline to clarify the correlation.
- Differentiate data points using markers.

Area Chart

Area Charts help show the part-to-whole relationships over time and demonstrate trends.

Best Practices

- Use contrasting colours for different areas.
- Be mindful of the stacking order.

Data Arrangement

Organising your data correctly in the spreadsheet is crucial. For example, time-series data should be in chronological order for Line Charts, and categorical data should be grouped for Column or Bar Charts.

Customisation and Advanced Features

Excel offers advanced customisation options:

1. **Labels:** Add data labels, axis titles, and legends.
2. **Trendlines:** Useful for Line and Scatter Plots.
3. **Error Bars:** For representing data variability.
4. **3D Effects:** Though flashy, they can sometimes reduce readability.

Tools and Add-ins for Advanced Visualisation

Excel supports various add-ins for more complex data visualisation, such as Power Map and other third-party tools.

PivotCharts

PivotCharts allows you to create charts from PivotTables, making it easier to analyse large data sets.

4.2 Formatting Charts in Excel

Formatting charts in Excel is a crucial step in data visualisation. It transforms a basic chart into a visually compelling representation of your data, making it easier for your audience to understand and interpret.

Importance of Formatting

Before diving into the technicalities, let's understand why formatting is necessary in data visualisation. Imagine you have a chart filled with data points, lines, and labels, but it looks cluttered and confusing. Your audience may struggle to extract meaningful insights from it. This is where formatting comes to the rescue. By applying formatting techniques, you can:

1. **Enhance Readability:** Proper formatting ensures that your chart is easy to read. It allows you to adjust font sizes, colours, and styles to make labels and text legible.
2. **Highlight Key Data:** Formatting allows you to emphasise specific data points, series, or categories. You can use colours, borders, and effects to draw attention to what matters most.
3. **Improve Aesthetics:** A well-formatted chart is visually appealing. It can make your presentation or report more engaging and professional.
4. **Tell a Story:** Formatting helps you convey your message effectively. You can use elements like chart titles, axis labels, and data labels to guide your audience through the data.

Key Formatting Elements in Excel Charts

Let's explore the critical formatting elements available in Excel charts and how to use them effectively:

1. **Chart Title:** The chart title is an essential component of your chart. It provides context and tells the viewer what the chart is about. To format the chart title:
 - Click on the chart title.
 - Use the "Format Chart Title" pane on the right to customise font, size, colour, and alignment.
2. **Axis Labels:** Axis labels are essential for understanding the chart's axes. You can format both the horizontal (X) and vertical (Y) axis labels:
 - Click on the axis labels.
 - Use the "Format Axis" pane to adjust font, size, colour, and alignment.
3. **Data Labels:** Data labels display the data values on the chart, making it easy to interpret. To format data labels:
 - Click on the data labels.
 - Use the "Format Data Labels" pane to control font, size, colour, and placement.
4. **Gridlines:** Gridlines are horizontal and vertical lines that help readers estimate values on the chart. You can format gridlines in various ways:
 - Right-click on gridlines and choose "Format Gridlines" to adjust line style, colour, and transparency.

5. **Chart Area:** The chart area is the entire chart, including all elements. You can format it to change the chart's background colour, border, and other properties:
 - Right-click anywhere within the chart area and choose "Format Chart Area."
6. **Legend:** The legend explains the colours or symbols used in the chart to represent different data series. To format the legend:
 - Click on the legend.
 - Use the "Format Legend" pane to modify font, size, colour, and position.
7. **Data Series:** Each data series in your chart can be formatted individually. This is useful when you want to emphasise specific data or make it stand out:
 - Click on the data series you want to format.
 - Use the "Format Data Series" pane to adjust fill colour, border colour, and other style attributes.
8. **Chart Elements:** Excel allows you to format various chart elements, such as lines, markers, and error bars, to suit your needs. You can access these options by selecting the specific element and right-clicking or using the "Format" tab in the Excel ribbon.

Advanced Formatting Techniques

Some of the advanced formatting techniques that can take your charts to the next level are:

1. **Colour Schemes:** Selecting an appropriate colour scheme can significantly impact the visual appeal of your chart. Excel provides predefined colour schemes, or you can create custom ones to match your brand or theme. Consider using contrasting colours to differentiate the data series.
2. **Chart Styles:** Excel offers a variety of chart styles, which are combinations of colours, fonts, and effects. You can quickly change the chart style from the "Chart Styles" gallery in the Excel ribbon. Experiment with different styles to find the one that best suits your data.
3. **3D Effects:** To make your chart pop, you can apply 3D effects such as shadows, bevels, and reflections. Be cautious with these effects, as overuse can make the chart look cluttered. You can access 3D formatting options in the "Format Chart Element" pane.
4. **Data Labels Customisation:** Customising data labels allows you to provide additional context or information. You can display data labels above or below data points, inside

pie chart slices, or in other creative ways. Excel offers several options for positioning and formatting data labels.

5. **Trendlines:** Trendlines help identify patterns and trends in your data. You can format trendlines to change their appearance, add labels, and adjust their precision. Trendlines are particularly useful in line and scatter charts.
6. **Background Images:** Sometimes, you may want to use background images in your charts. This can effectively create thematic charts or add context to geographic data. Excel allows you to insert and format pictures to fit the chart's context.

Tips for Effective Chart Formatting

Here are some tips to keep in mind when formatting charts in Excel:

1. **Keep it Simple:** Avoid excessive formatting that can distract from the data. Simplicity often leads to better clarity.
2. **Consistency:** Use consistent colours, fonts, and styles throughout your charts for a cohesive look.
3. **Accessibility:** Ensure your charts are accessible to all viewers, including those with visual impairments. Use high-contrast colours and provide alternative text where necessary.
4. **Review and Test:** Review your formatted chart to ensure it effectively conveys the intended message. Test it with different audiences to gather feedback.
5. **Context Matters:** Consider the context in which the chart will be viewed. Formatting may differ for a presentation, report, or web page.

SELF-ASSESSMENT QUESTIONS – 3

11. Which chart type is NOT suitable for showing data trends over time?

- a) Pie chart
- b) Line chart
- c) Area chart
- d) Column chart

12. Which chart type is NOT suitable for showing data trends over time?

- a) Pie chart
- b) Line chart
- c) Area chart
- d) Column chart

13. What is a 'Sparkline'?

- a) A tiny chart in a worksheet cell
- b) A line that sparks up when you click it
- c) A chart with spark effects
- d) A type of formula

14. Which chart type is best for comparing parts of a whole?

- a) Line chart
- b) Scatter chart
- c) Pie chart
- d) Column chart

15. How can you change the chart type after creating a chart?

- a) Delete and recreate the chart.
- b) Go to Chart Tools -> Design -> Change Chart Type
- c) Go to Insert -> Charts -> Change Type
- d) It's not possible to change the chart type.

5. SUMMARY

In this comprehensive unit on Excel in Data Visualization, we have traversed from the foundational elements to advanced functionalities that Excel offers. We started with an introduction to Microsoft Excel, emphasising its importance as a versatile data management and visualisation tool. Understanding the basics of spreadsheets was our first milestone, where you learned about rows, columns, cells, and the essence of data organisation.

From there, we delved into the nuances of data types and their formatting options. Knowing how Excel recognises numbers, text, dates, and even custom types offers a solid foundation for accurate data visualisation. We also explored the numerous formatting options, stressing the impact of aesthetics and readability on effective data presentation.

The unit then elevated to the core topic of visualisation capabilities. A range of charts and graphs were discussed, each with its best practices and ideal use cases. Excel offers various options to bring your data to life, from column and line charts for trend analysis to pie charts for part-to-whole relationships. Customisation techniques and advanced features were also covered to ensure you're equipped with the tools to create compelling, insightful visuals.

For those new to programming, the ease and versatility of Excel serve as an entry point into the data-driven decision-making process. The skills you've acquired are not just theoretical but have immense practical applications in various professional settings.

6. GLOSSARY

1. **Web Scraping:** A method used to extract large amounts of data from websites where data is extracted and saved to a local file in your computer or a database in table (tabular) format.
2. **BeautifulSoup:** A Python library used for web scraping to pull the data from HTML and XML files.
3. **rvest:** A simple web scraping package for R, inspired by libraries like BeautifulSoup.
4. **Pandas:** A software library written for the Python programming language for data manipulation and analysis. It is beneficial for handling large data structures like data frames.
5. **dplyr:** A grammar of data manipulation in R, providing tools for efficiently manipulating datasets.
6. **Data Cleaning:** The process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted.
7. **Data Wrangling/Munging:** The cleaning, structuring and enriching of raw data into a desired format for better decision-making.
8. **Job Titles:** A term that describes the responsibilities of a position. It also defines the rank or level of the position.
9. **Job Descriptions:** A general or detailed written statement of what a job entails, including the job title, job location, job summary, duties, responsibilities, and skills needed to perform the job.
10. **Job Requirements:** The specific qualifications and skills needed to perform a job.
11. **Data Visualisation:** The graphical representation of information and data, typically using statistical graphics, plots, information graphics and other tools.
12. **matplotlib:** A plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications.
13. **Seaborn:** A Python data visualisation library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

14. **ggplot2:** A data visualisation package for the R programming language. It is based on the grammar of graphics, the idea that you can build every graph from the same few components: a data set, a set of geometries, and a coordinate system.
15. **Job Market Trends:** The patterns and trends occurring in the job market in response to economic, technological, and social changes.
16. **HTML:** HyperText Markup Language is the standard markup language for documents designed to be displayed in a web browser.
17. **CSS Selector:** A pattern used to select the element(s) you want to style or manipulate in some way. In web scraping, CSS selectors are used to identify the data on the webpage we want to extract.
18. **Web Elements:** HTML elements that define the structure of a webpage and contain the webpage's content. They include headings, paragraphs, links, images, etc.
19. **API (Application Programming Interface):** A set of rules that allows programs to talk to each other, defining the methods and data formats a program can use to communicate with other programs.
20. **Data Frame:** A two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns. Pandas DataFrame consists of three principal components: the data, rows, and columns.
21. **URL (Uniform Resource Locator):** The address used to find and retrieve any data or information (pages, files, images, or programs) on the web.
22. **HTTP Requests:** A request message from a client to a server over the internet or other networks. HTTP (Hypertext Transfer Protocol), the protocol the World Wide Web uses, defines the syntax and semantics the client and server must adhere to.
23. **HTTP Response Codes:** Standard codes given by website servers on the internet. The codes help identify the cause of the problem when a web page or other resource does not load properly.
24. **JSON (JavaScript Object Notation):** A standard text-based format for representing structured data based on JavaScript object syntax, commonly used for describing and transmitting data on websites.
25. **Data Mining:** The process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.

26. **Data Analysis:** Inspecting, cleansing, transforming, and modelling data to discover useful information, inform conclusions, and support decision-making.
27. **CSV (Comma Separated Values):** A simple file format used to store tabular data, such as a spreadsheet or database. Each line of the file is a data record, and each record consists of one or more fields separated by commas.
28. **XPath (XML Path Language):** A language that uses a browsing paradigm for XML documents (used for selecting nodes from an XML document), which is also useful when the user doesn't have an "id" or "class", and the user wants to extract the data.
29. **Robots.txt:** A text file webmasters create to instruct web robots (typically search engine robots) how to crawl pages on their website.
30. **Web Crawler:** An internet bot that systematically browses the World Wide Web, typically for Web indexing (web spidering).
31. **Text Mining:** The process of deriving high-quality information from text. It involves the discovery by computer of new, previously unknown information by automatically extracting information from different written resources.
32. **Unstructured Data:** Information that either does not have a predefined data model or is not organised in a predefined manner. Unstructured information is typically text-heavy but may contain data such as dates, numbers, and facts.
33. **Structured Data:** Data that has been organised into a formatted repository, typically a database, so that its elements can be made addressable for more effective processing and analysis.
34. **Regular Expression (RegEx):** A sequence of characters that forms a search pattern, mainly used in pattern matching with strings or string matching.
35. **Web Page Parsing:** The process of analysing a string of symbols, either in natural language, computer languages or data structures, conforming to the rules of a formal grammar.
36. **Selenium:** A powerful tool for controlling a web browser through the program. It is functional for all browsers, works on all major OS and its scripts are written in various languages, i.e. Python, Java, C#, etc.
37. **Scrapy:** An open-source and collaborative web crawling framework for Python. It is used to extract the data from the web page with the help of selectors based on XPath.

38. **Web Driver:** A collection of open-source APIs used to automate a web application's testing. Description: Selenium WebDriver is a collection of open web APIs that can automate the execution of a web application on various browsers and operating systems.
39. **Data Set:** A collection of related sets of information composed of separate elements that a computer can manipulate as a unit.
40. **Data Collection:** The process of gathering and measuring information on targeted variables in an established system enables one to answer relevant questions and evaluate outcomes.

7. TERMINAL QUESTIONS

1. Explain the importance of understanding data types in Excel for effective data visualisation.
2. Describe how spreadsheet organisation can impact the ease and effectiveness of creating visualisations in Excel.
3. Discuss three types of charts that are most effective for showing trends over time and justify your choices.
4. What are some essential formatting techniques to enhance the readability and interpretability of Excel charts?
5. Outline the steps involved in creating a pie chart in Excel and discuss when it is most appropriate.
6. Describe the role of Excel's 'Pivot Table' feature in data visualisation.
7. Compare and contrast the uses of column charts and line charts in Excel.
8. How can conditional formatting be utilised for data visualisation in Excel?
9. Explain the significance of data formatting in Excel and its impact on visual representation.
10. Discuss the main challenges a working professional might face when new to Excel and how they can overcome them.
11. Enumerate the steps to change the chart title and axis labels in Excel.
12. Why is it essential to evaluate the effectiveness of various Excel chart types in representing different data sets?

13. Describe the basic building blocks of an Excel spreadsheet and their roles.
14. How can Excel's sorting and filtering features aid in data visualisation?
15. Discuss the use cases where Excel for data visualisation is preferable over more specialised software like Tableau.
16. Explain how to customise the colour scheme of a chart in Excel.
17. What are some of the limitations of using Excel for data visualisation, and how can they be mitigated?
18. How does understanding the formatting options in Excel contribute to better data visualisation?
19. Describe the steps to create an Excel bar graph and explain when to use it.
20. Discuss how the knowledge of Excel can give you a competitive edge in today's data-driven world.

8. ANSWERS

SAQ Answers:

1. (Answer: B. An intersection between a row and a column)
2. (Answer: B. A)
3. (Answer: B. A1:C3)
4. (Answer: C. Swiping on the touchpad)
5. (Answer: B. Select the entire column)
6. (Answer: B. It formats cells based on specific conditions.)
7. (Answer: B. CONCATENATE)
8. (Answer: B. It wraps text within a cell)
9. (Answer: C. It combines multiple cells into a single cell)
10. (Answer: C. Using the 'Fill Color' option)
11. (Answer: A. Pie chart)
12. (Answer: B. Insert -> Charts -> Chart Type)
13. (Answer: A. A tiny chart in a worksheet cell)
14. (Answer: C. Pie chart)
15. (Answer: B. Go to Chart Tools -> Design -> Change Chart Type)

Terminal Questions Answers:

1. Refer Section: 3.1
2. Refer Section: 2.2
3. Refer Section: 4.1
4. Refer Section: 4.2
5. Refer Section: 4.1
6. Refer Section: 4.1
7. Refer Section: 4.1
8. Refer Section: 3.2
9. Refer Section: 3.2
10. Refer Section: 2
11. Refer Section: 4.2
12. Refer Section: 2
13. Refer Section: 2.2
14. Refer Section: 3.2
15. Refer Section: 4.1
16. Refer Section: 4.2
17. Refer Section: 4.1
18. Refer Section: 3.2
19. Refer Section: 4.1
20. Refer Section: 2

9. SUGGESTED BOOKS AND E-REFERENCES

1. "Spreadsheet Software Programs | Excel Free Trial." Microsoft Excel, Microsoft.
2. "How to use Excel – Tutorial for beginners." Digital Unite, Digital Unite.
3. Walkenbach, John. "Excel 2019 Bible." Wiley, 2018.
4. "Data Types in Excel." Microsoft Support, Microsoft.
5. "Advanced Excel Data Types." ExcelJet, ExcelJet.
6. Winston, Wayne. "Microsoft Excel 2019 Data Analysis and Business Modeling." Microsoft Press, 2019.
7. "Excel Data Types: A Comprehensive Guide." AbleBits, AbleBits.

