SUMMARY: INTRODUCTION TO MS EXCEL-I

SESSION OVERVIEW:

By the end of this session, students will be able to-

- Understand how MS Excel helps in Data Analytics.
- Understand the interface of MS Excel.
- Perform different operations using MS Excel.

KEY TOPICS AND EXAMPLES:

1. Understand how MS Excel helps in Data Analytics:

Microsoft Excel is a popular choice among users for carrying out different analytical activities since it provides a number of benefits for data analytics. Using Microsoft Excel for data analytics has the following main benefits:

- **Ease of use:**
  Excel is renowned for its intuitive features and user-friendly layout, which enable users with different degrees of technical expertise to utilize it with ease. Basic data analysis functions like sorting, filtering, and charting are simple for users to do without a lot of experience.
- **Familiarity:**
  Excel is a ubiquitous tool that many users are already familiar with due to its widespread use in businesses, academia, and personal computing. This familiarity reduces the learning curve for users transitioning to data analytics tasks in Excel, as they can leverage their existing knowledge of the software.
- **Versatility:**
  Excel provides a multitude of features to meet various data analytics requirements. In addition to writing custom macros using Visual Basic for Applications (VBA) to automate repetitive activities, users may create interactive dashboards, perform basic statistical analysis, and create pivot tables and pivot charts.
- **Data Import and Export:** Excel supports importing data from various sources such as databases, text files, CSV files, and online sources. Users can easily import data into Excel for analysis and export the results to other formats for sharing or further analysis.
- **Data Cleaning and Transformation:** Excel provides tools for cleaning and transforming data, such as removing duplicates, handling missing values, and formatting data. Users can use built-in functions and formulas to preprocess data before analysis, ensuring data quality and integrity.
- **Data Visualization:** Excel offers a variety of chart types and visualization options for creating visually appealing and informative charts and graphs. Users can easily create bar charts, line charts, pie charts, scatter plots, and more to visualize their data and communicate insights effectively.
- **Integration with Other Tools:** Excel integrates seamlessly with other Microsoft Office applications, allowing users to share data and reports between Excel, Word, PowerPoint, and Outlook. Users can also connect Excel to external data sources and analysis tools for more advanced analysis.

- **Cost- effectiveness**: Excel is already available as a part of the Microsoft Office Suite, which makes it a cost-effective option compared to specialized data analysis software.

Industry importance of MS Excel: (10-15 min)

*(**Comment for instructor**: Discuss other industries if required.)*

Almost all companies use excel because the higher management is familiar with the tool and do not require complex IT setup to access data. Also, it is helpful for rapid prototyping for quick decision making, and excel is the common language across all industries. Mentioned below are some specific industry examples which rely on Excel.

- **Banking and finance:**
  Excel is extensively used in finance and banking for financial modeling, budgeting, forecasting, risk analysis, and reporting. Financial analysts, accountants, and investment professionals rely on Excel to analyze financial data, create complex financial models, and generate reports for decision-making and regulatory compliance.
- **Business and consulting:**
  Excel is a staple tool in business and consulting firms for data analysis, market research, project management, and strategic planning. Consultants use Excel to analyze business data, develop financial models, conduct feasibility studies, and create presentations for clients.
- **Manufacturing and engineering:**
  Excel is widely used in manufacturing and engineering industries for inventory management, production planning, quality control, and process optimization. Engineers and operations managers use Excel to analyze production data, track inventory levels, and optimize manufacturing processes for efficiency and cost savings.
- **Healthcare and pharmaceutical:**
  Excel is used in healthcare and pharmaceutical industries for patient data management, clinical trials, drug development, and medical research. Healthcare professionals use Excel to organize patient records, analyze medical data, and track clinical outcomes for research and decision-making purposes.
- **Retail and e-commerce:**
  Excel is used in retail and e-commerce industries for inventory management, sales analysis, customer segmentation, and marketing campaigns. Retailers use Excel to track sales data, analyze customer behavior, and forecast demand for products and services.

2. Understand the interface of MS Excel:

2. A. How to open MS Excel: (5 mins)

- Click on the search bar present below in the screen.
- Type MS Excel in the search bar.
- Click on MS Excel.
- A screen opens up as shown in the picture below. *(Figure 1)*
- Click on the blank workbook and a new workbook will open. *(Figure 2)*

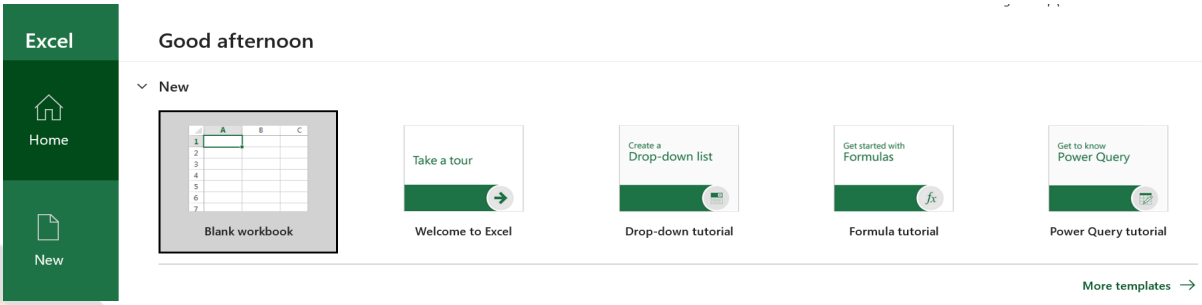


Figure 1: Figure showing steps related to opening of workbook

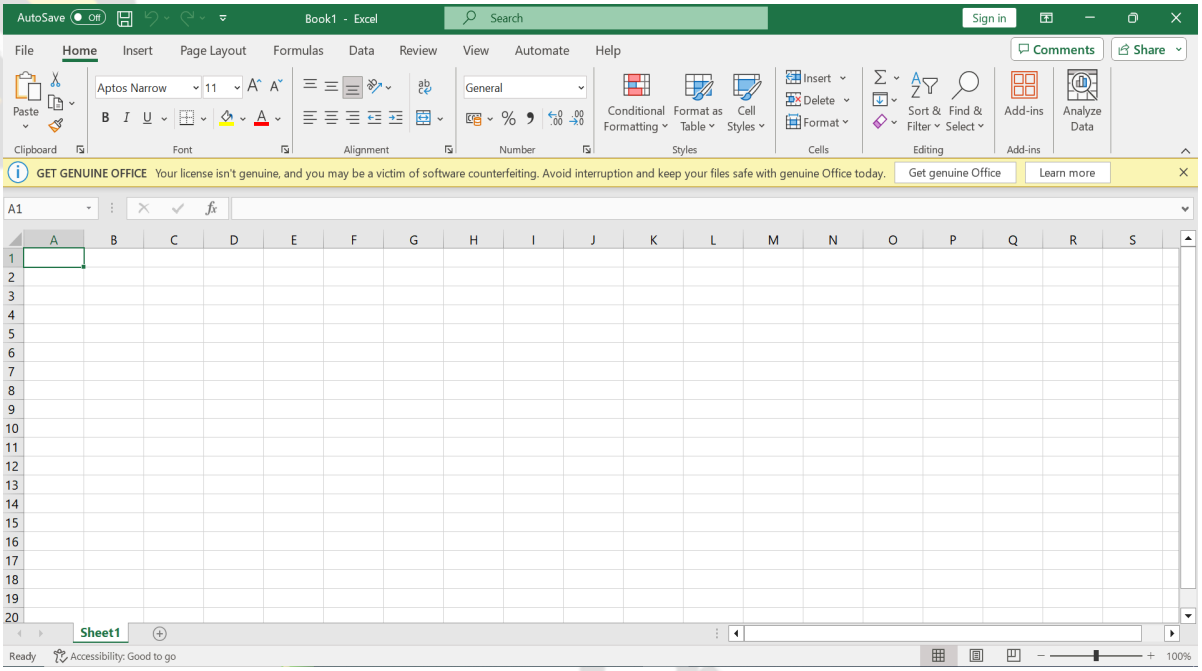


Figure 2: Shows the picturization of a workbook

2. B. Introduction to the Excel ribbon: (10 mins)

The idea is to introduce ribbon in excel to the students.
In Microsoft Excel, the Ribbon is organized into tabs and groups, each containing related commands, functions, and tools. Here's an explanation of the tabs and groups in MS Excel:

Tabs:

- The Ribbon is divided into tabs, which are displayed horizontally across the top of the Excel window.
- Each tab represents a different set of commands and tools related to specific tasks or functions in Excel.
- The default tabs in Excel include: Home, Insert, Page Layout, Formulas, Data, Review, and View. Additional tabs may appear dynamically depending on the context of your current task, such as Chart Tools or Picture Tools.

Groups:

- Within each tab, commands are further organized into groups, which group related commands together for easier access.

- Groups are displayed as sections within the tab, often separated by lines or space.
- Each group contains a set of related commands or tools that are commonly used together for specific tasks.
- For example, within the "Home" tab, common groups include Clipboard, Font, Alignment, Number, Styles, Cells, and Editing. Each group contains commands related to formatting, aligning, and editing cells.

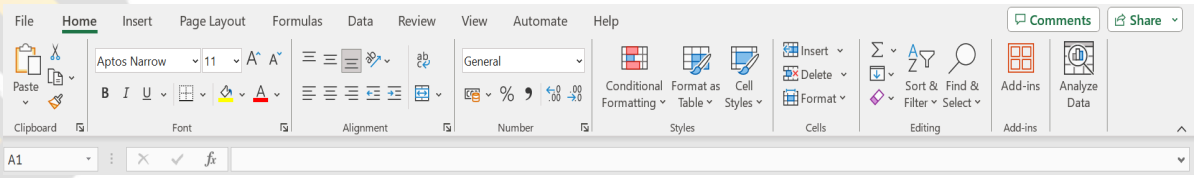


***Figure 3**: Ribbon associated to home tab*

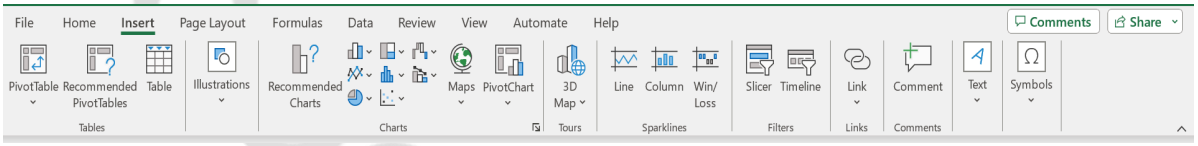


***Figure 4**: Ribbon associated to insert tab*

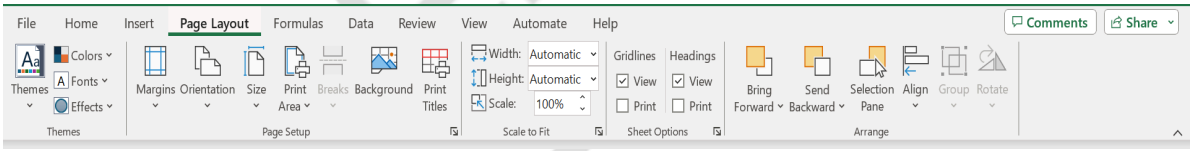


***Figure 5:** Ribbon associated to page layout tab*

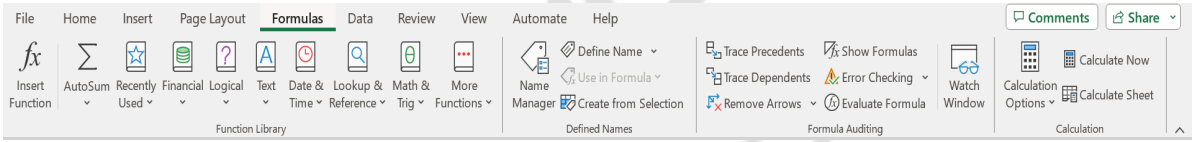


***Figure 6**: Ribbon associated to formulas tab*

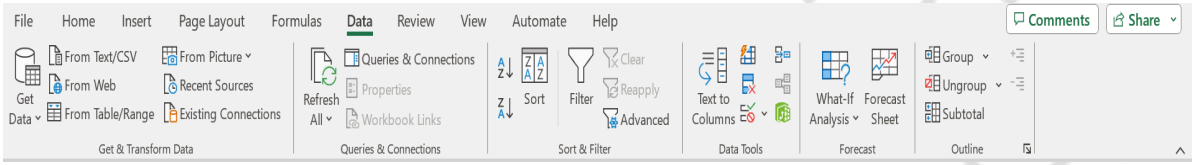


***Figure 7**: Ribbon associated to data tab*

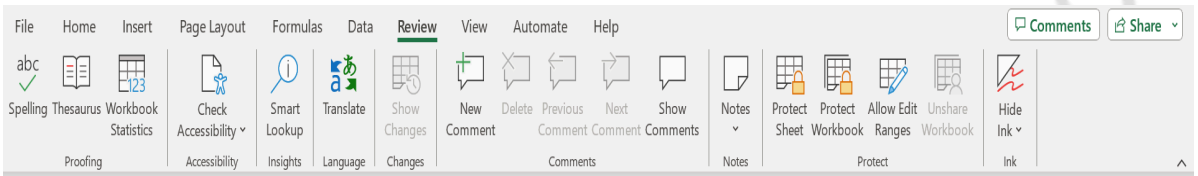


***Figure 8:** Ribbon associated to Review tab*

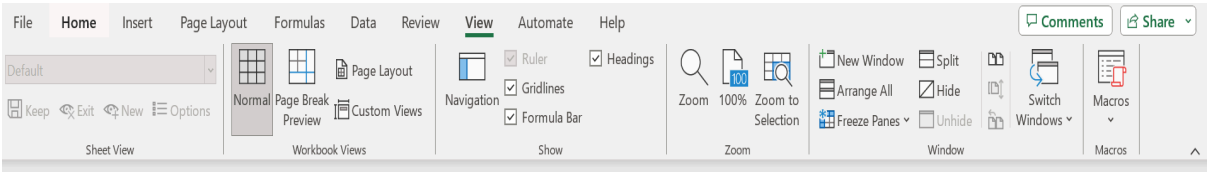


***Figure 9**: Ribbon associated to View tab*

2. C. Introduction to different parts of excel: (5mins)

Rows:

- Rows in Excel are horizontal lines of cells that are identified by numbers along the left side of the worksheet, typically starting with "1" at the top.
- Each row extends from the left edge of the worksheet to the right edge, and contains a series of cells.
- Rows are numbered sequentially, with the first row at the top of the worksheet being row 1, the second row below it being row 2, and so on.
- Rows provide a way to organize and structure data horizontally in Excel.

Columns:

- Columns in Excel are vertical lines of cells that are identified by letters along the top of the worksheet, typically starting with "A" at the leftmost column.
- Each column extends from the top edge of the worksheet to the bottom edge, and contains a series of cells.
- Columns are lettered alphabetically, with the first column at the left edge of the worksheet being column A, the second column next to it being column B, and so on.
- Columns provide a way to organize and structure data vertically in Excel.

Cells:

- Cells are the individual rectangular boxes formed by the intersection of rows and columns in Excel.
- Each cell is identified by a unique combination of its row number and column letter. For example, the cell at the intersection of row 10 and column E is referred to as cell "E10." ***(Figure 10)***
- Cells are the basic units of data storage in Excel, and can contain various types of data such as numbers, text, formulas, and dates.
- Cells can be formatted, edited, and manipulated individually or in groups, making them the building blocks for creating and organizing data in Excel.

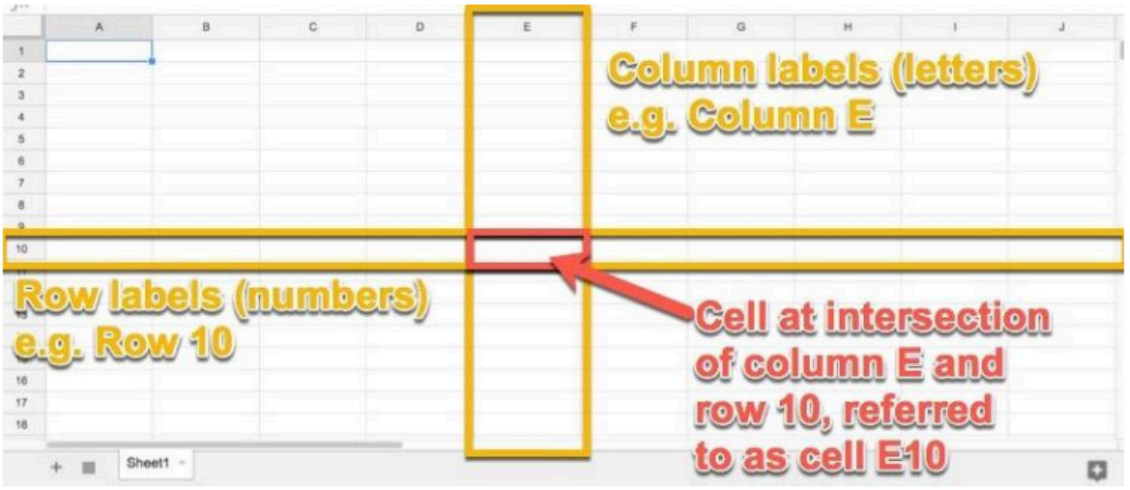
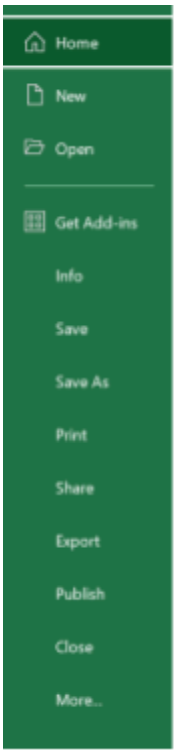


Figure 10: *Representation of rows, columns and cell in MS Excel*

2. D. How to save data in excel:(5 mins)

*(**Comment for instructor**: Before introducing these terms practically, describe briefly the uses of each of the terms mentioned below)*



New: This will help to create new workbook for you.

- Searching for the type of template you want, using keywords that describes it, such as "calendar" or "invoice".
- Clicking suggested search located below the search field.
- Select the template you want, if it already appears on the page.

Open: Using open, you can open an already existing file in excel. Once, you click on the tab, a page appears where you can choose the file you desire for.

Save or save as: If you're saving the workbook for the first time or want to save it with a different name/location, click on "Save As." If you've already saved the workbook and want to overwrite the existing file with your changes, simply click on "Save."

Print: Open your Excel worksheet. Go to the "File" tab. Select "Print." Adjust settings like page layout, print area, and number of copies. Click "Print" to send the worksheet to the printer.

Export: Open your Excel worksheet. Select the data you want to export.Go to the "File" tab. Choose "Save As" or "Export." Select the file format (e.g., CSV, PDF). Choose the location to save the file. Click "Save" or "Export" to save the data in the chosen format.

Figure 11*: Representation of save, save as, open, print and import in MS Excel*

Important: At times, students tend to save Excel files in CSV format. Saving files in CSV format will lead to the following problems:

- ***Loss of Formatting:*** CSV files do not support Excel's formatting features such as font styles, colors, cell merging, or conditional formatting. When an Excel file is saved as a CSV, all formatting information is stripped away, resulting in plain text data without any visual representation.
- ***Loss of Multiple Worksheets:*** Excel workbooks can contain multiple worksheets, each with its own set of data. However, CSV files can only represent a single table of data. When you save an Excel workbook with multiple worksheets as a CSV, only one worksheet will be preserved, leading to loss of data from other sheets.
- ***Loss of Cell Types***: Excel supports various cell types such as numbers, text, dates, formulas, and hyperlinks. However, CSV files can only store plain text data. As a result, when you save an Excel file as a CSV, any non-text data, such as formulas or hyperlinks, will be converted to plain text, potentially resulting in loss of functionality or accuracy.

2. E. How to import data in MS Excel: (5 mins)

Open Excel and create a new worksheet.

- Go to the "Data" tab.
- Click on "Get Data" or "From Text/CSV" depending on your Excel version.
- Choose the source of your data (e.g., text file, database, web).
- Follow the prompts to select the file or connect to the data source.
- Adjust import settings if needed (e.g., delimiter, encoding).
- Click "Load" or "Import" to bring the data into Excel.

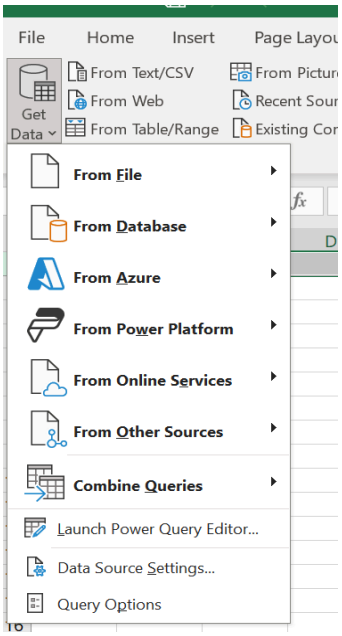


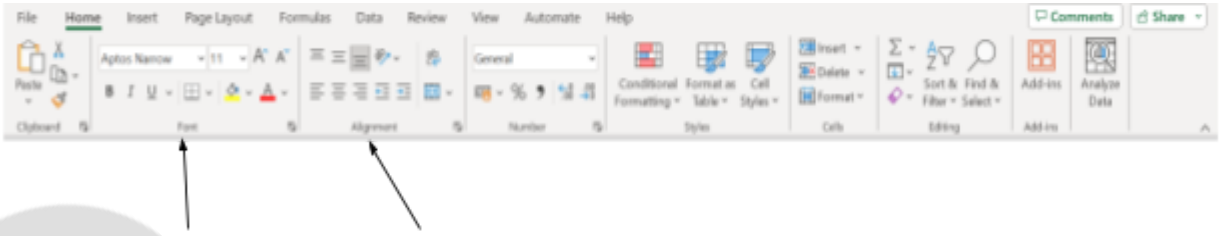
Figure 12*: Represents different sources from which data can be imported*

3. Basics of MS Excel:

3. A. Shortcuts in Excel:

Windows key combination	Mac key combination	Uses of the key
Arrow Keys	Arrow Keys	Move one cell up, down, left, or right
Tabs	Tabs	Move one cell to the right
Shift + Tabs	Shift + Tabs	Move one cell to the left
Ctrl + End	Control-End *or* *Control-Fn-right arrow*	Go to last cell in data region
Home	Home *or* Fn-left arrow	Move to the beginning of the row
Ctrl + Home	Control-Home *or* *Control-Fn-Left arrow*	Move to the beginning of the worksheet
Ctrl + Arrow Left / Ctrl + Arrow Right	Cmd-left arrow/ Cmd-right arrow	Go to left-most / right-most cell in current row
Ctrl + Arrow Down/ Ctrl + Arrow Up	Cmd-down arrow/ Cmd-up arrow	Go to bottom-most / to-most cell in current column
Ctrl + Page Down / Ctrl + Page Up	Control-Page Down/ Control-Page Up	Move to next / move to previous worksheet
Alt + Page Down / Alt + Page Up	Option-Page Up/ Option-Page Down *or* *Fn-Option-up arrow /Fn-Option-down arrow*	Move one screen to the right / one screen to the left
Ctrl + Shift + Page Up/Page Down	Cmd-shift-[/ Cmd-shift-]	Switch sheets

3. B. Basic formatting in MS Excel:

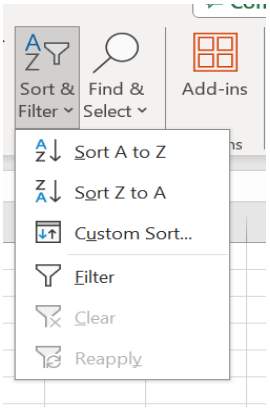


1. Borders
2. Bold (CTRL+B)
3. Underline (CTRL+U)
4. Text color
5. Cell color

4. Perform different functions, operations and formulas using MS Excel:

4. A. Basic operations: (30 mins)

- **Sorting:**
  - **Uses:**
    - Sorting in Excel allows users to:
    - Organize data alphabetically, numerically, or by custom criteria.
    - Quickly find specific data points or patterns.
    - Analyze data more efficiently.
    - Prepare data for presentation or reporting.
    - Identify trends or outliers.
    - Compare and contrast different data sets.
    - Improve data visualization and readability.
  - **Steps:**
    - Select the column you want to sort by clicking on the column header. Then, click on the "Sort A to Z" button in the Ribbon to sort the data in ascending order or the "Sort Z to A" button to sort in descending order.
    - If you have multiple columns and want to sort by specific criteria, you can click on the "Sort & Filter" button in the Ribbon and choose "Custom Sort..." from the drop-down menu. This will open the Sort dialog box where you can specify your sorting criteria.



- **Freezing:**
    - **Uses:**
        - Keep headers or labels visible while scrolling.
        - Maintain a reference point for analysis.
        - Compare data across different sections.
        - Ensure consistent formatting during navigation.
        - Simplify data entry and editing processes.
        - Improve readability and analysis of large datasets.
        - Enhance efficiency in data manipulation and reporting.
    - **Steps:**
        - Click on the cell below the row you want to freeze, or to the right of the column you want to freeze.
        - Go to the "View" tab.
        - Click on "Freeze Panes" in the "Window" group.
        - Choose either "Freeze Top Row" to freeze rows or "Freeze First Column" to freeze columns.

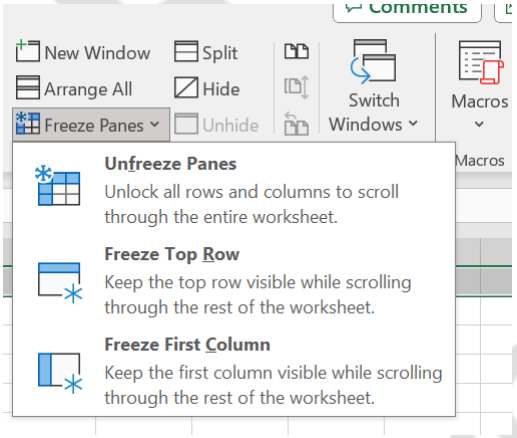


Figure 14: *Represents freeze feature*

- **Find and replace:**
    - **Uses:**
        - Locating specific values or text.
        - Editing multiple instances of a value or text.
        - Cleaning up data inconsistencies or errors.
        - Updating formatting attributes.
        - Correcting spelling mistakes or typos.
        - Transforming data by replacing patterns or text strings.
        - Performing conditional replacements based on specific criteria.
    - **Steps:**
        - Press Ctrl + F to open the Find and Replace dialog.
        - Enter the text you want to find in the "Find what" field.
        - Optionally, enter replacement text in the "Replace with" field.
        - Click "Find Next" to locate the first instance, or "Replace" to replace it.
        - Use "Replace All" to replace all instances at once.

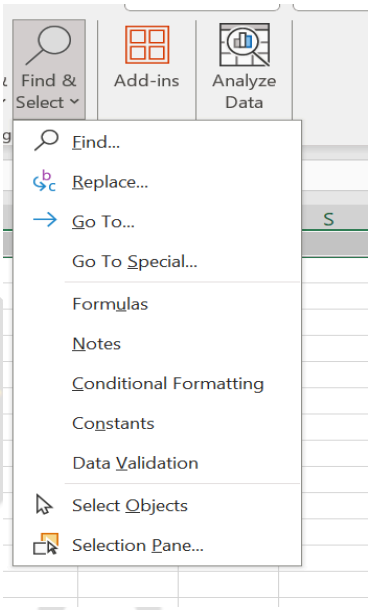


Figure 15: Represents find and replace in Excel

- **Paste special:**
    - **Uses:**
        - Paste data with specific formatting.
        - Perform mathematical operations during pasting.
        - Copy formulas without adjusting cell references.
        - Paste data as values, removing formulas or formatting.
        - Transpose data, switching rows to columns or vice versa.
        - Paste data as links or pictures for dynamic updates.
        - Preserve data integrity and structure during copying and pasting tasks.
    - **Different types of paste special:**
        - **Values:** Paste only the values of the copied cells, without formatting or formulas.
        - **Formulas:** Paste the formulas from the copied cells, adjusting cell references as needed.
        - **Formats:** Paste only the formatting of the copied cells, such as font style, color, and borders.
        - **Values and Number Formats:** Paste both the values and the number formats of the copied cells.
        - **Transpose:** Transpose the copied data, switching rows to columns or columns to rows.
    - **Steps:**
        - Copy the data.
        - Right-click the destination cell.
        - Choose your "Paste Special" from the menu.
        - Select the desired option.
        - Click OK.

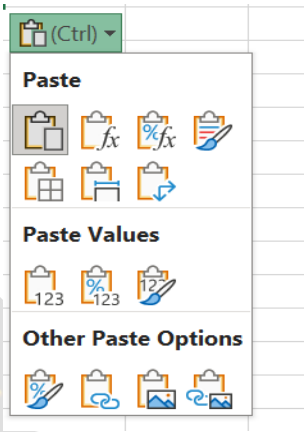


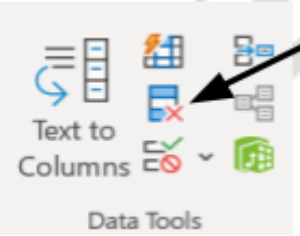
Figure 16: Represents different paste special features and transpose functionality in MS Excel

- **Transpose functionality:**
  - **Uses:**
    - Data reorganization for analysis.
    - Creating summary tables.
    - Adjusting data layout for formatting.
    - Consolidating multiple datasets.
    - Preparing data for charting.
    - Importing/exporting data from/to other applications.
  - **Steps:**
    - Select the range of cells you want to transpose.
    - Right-click and choose "Copy" or press Ctrl + C to copy the data.
    - Right-click on the cell where you want to paste the transposed data.
    - Choose "Paste Special" from the context menu.
    - Check the "Transpose" option.
    - Click "OK" to apply the transpose operation.

4. B. Important functionalities:

- **Filter operation:**
  - **Uses:**
    - View specific data subsets by hiding irrelevant rows based on specified criteria.
    - Analyze data more efficiently by focusing on relevant information.
    - Identify trends, patterns, or outliers within datasets.
    - Sort data alphabetically, numerically, or by custom criteria.
    - Apply multiple filters simultaneously to refine data views further.
    -
  - **Steps:**
    - Select the data range.
    - Go to the "Data" tab.
    - Click on the "Filter" button.
    - Use the filter arrows in the column headers to select criteria.

        - Excel will filter the data based on your selection, showing only the rows that meet the criteria.
        - To remove the filter, click on the filter arrow again and select "Clear Filter."
- **Deleting duplicates:**
    - **Uses:**
        - Ensures accuracy
        - Improve statistical analysis
        - Enhance data quality
        - Reduce bias
        - Streamline processing
        - Facilitate insights
        - Support decision-making.
    - **Steps:**
        - Select the range of cells containing your data.
        - Go to the "Data" tab.
        - Click on "Remove Duplicates" in the Data Tools group.
        - Choose the columns where you want to find duplicates.
        - Click OK.
        - Excel will remove duplicate rows based on the selected columns, keeping only unique values.
        - Ensure to review the operation as it's irreversible.



Datasets: The instructor has performed all the operations on this Dataset throughout this module.