# Introduction to Spreadsheets

Reading data into Excel

Definition: This involves methods for importing data into an Excel worksheet so it can be stored, managed, and analyzed. You can import from various formats, including text files (.txt, .csv) and other Excel workbooks.

Example:

* From a CSV file: To import customer sales data from a sales.csv file, go to the Data tab > Get & Transform Data > From Text/CSV. Select the file, and Excel will load a preview. You can then specify the delimiter (e.g., a comma) before loading the data into a new worksheet or an existing one.

Definition: The process of cleaning, organizing, and transforming raw data to prepare it for analysis. It involves several key techniques.

Examples of common techniques:

* Sorting: Arranging data in a specific order (ascending or descending) based on the values in one or more columns. You can sort a sales list alphabetically by product name.
* Removing duplicates: Identifying and deleting duplicate entries to ensure data integrity. This is often used to create a clean list of unique customer emails.
* Text to Columns: Splitting text from one column into multiple columns based on a delimiter. For instance, separating a full name from a single column into "First Name" and "Last Name" columns.
* Filtering: Temporarily hiding irrelevant rows to focus on specific subsets of data. For example, viewing only sales from a specific region.

Basic functions in Excel

Definition: Predefined formulas that perform specific calculations on your data. Mastering a few fundamental functions is essential for basic data analysis.

Examples:

* =SUM(A1:A10): Adds all the numerical values in the range of cells from A1 to A10.
* =AVERAGE(B1:B20): Calculates the average (mean) of all numbers in the specified range.
* =COUNT(C:C): Counts the number of cells in column C that contain a numerical value.
* =MAX(D:D) and =MIN(D:D): Finds the highest and lowest numerical value, respectively, in column D.

Spreadsheet functions to organize data

IF function

Definition: A logical function that checks a condition and returns one value if the condition is TRUE and another if FALSE.

Syntax: =IF(logical\_test, value\_if\_true, value\_if\_false)

Example:

* Scenario: You have a list of sales figures in column B and want to flag transactions over $500 as "High".
* Formula: =IF(B2>500,"High Sales","Low Sales").

Nested IF

Definition: The practice of placing one or more IF functions inside another to test multiple criteria sequentially and return different outcomes.

Example:

* Scenario: Assigning a division based on student marks in cell B2.
* Formula: =IF(B2>=60,"1st Division",IF(B2>=30,"2nd Division",IF(B2>=10,"3rd Division","Fail")))
* Breakdown: The formula first checks if the score is 60 or higher. If false, it checks if it's 30 or higher, and so on. This allows for multiple outcomes.

VLOOKUP (Vertical Lookup)

Definition: Searches for a value in the first column of a range (or "table array") and returns a corresponding value from a specified column in the same row.

Syntax: =VLOOKUP(lookup\_value, table\_array, col\_index\_num, [range\_lookup])

Example:

* Scenario: You have a list of employee IDs in one sheet and need to find the department for a specific employee ID from another sheet.
* Formula: =VLOOKUP(103, A1:C5, 3, FALSE) where:
  + 103 is the lookup\_value (Employee ID).
  + A1:C5 is the table\_array containing the employee data.
  + 3 is the col\_index\_num for the Department column.
  + FALSE specifies an exact match for the Employee ID.

HLOOKUP (Horizontal Lookup)

Definition: Similar to VLOOKUP, but it searches for a value in the *top row* of a table and returns a value from a specified row below it.

Syntax: =HLOOKUP(lookup\_value, table\_array, row\_index\_num, [range\_lookup])

Example:

* Scenario: You have monthly sales data with headers like "Jan," "Feb," "Mar" in the top row and want to find the sales figure for "Feb" from the "Sales" row.
* Formula: =HLOOKUP("Feb", B1:D4, 2, FALSE) where:
  + "Feb" is the lookup\_value.
  + B1:D4 is the table\_array covering the months and sales figures.
  + 2 is the row\_index\_num for the "Sales" row.
  + FALSE specifies an exact match.

Introduction to filtering, pivot tables, and charts

Data filtering in Excel

Definition: A quick way to selectively display data that meets specific criteria while temporarily hiding irrelevant rows.

Implementation:

1. Select a cell within your data range.
2. Navigate to the Data tab and click the Filter button.
3. Click the dropdown arrow that appears in the header of the column you want to filter.
4. Deselect (Select All) and choose the specific values you want to display. Alternatively, use options like Number Filters (e.g., Greater Than) or Text Filters.
5. Example: To see only sales from the "East" region, you would filter the "Region" column and select only "East".

Use of pivot tables in Excel

Definition: An interactive tool for summarizing and reorganizing large datasets to identify patterns and trends. It allows you to drag and drop fields to dynamically change your data's perspective.

Implementation:

1. Click any cell inside your dataset.
2. Go to the Insert tab and click PivotTable.
3. Confirm the data range and choose where to place the new pivot table (e.g., a new worksheet).
4. Use the PivotTable Fields pane to arrange fields into four areas:
   * Rows: Categories to display along the left side.
   * Columns: Categories to display across the top.
   * Values: The numerical data to summarize (e.g., sum or average of sales).
   * Filters: Fields to filter the entire table by.
5. Example: A sales manager could use a pivot table to see the total sales for each product category per month. They would drag "Product Category" to Rows, "Month" to Columns, and "Sales" to Values.

Introduction to charts in Excel

Definition: Graphical representations of data that make it easier to interpret and analyze patterns, trends, and comparisons.

Implementation:

1. Select the data you want to visualize.
2. Go to the Insert tab and click Recommended Charts or choose a specific chart type (e.g., Column, Line, or Pie).
3. Excel will display a preview. Select the desired chart and click OK to insert it.
4. Example: A marketing team wants to visualize monthly sales. They can select the "Month" and "Sales" columns and create a line chart to see the sales trend over time.

# OR

**Introduction to spreadsheets**

* Spreadsheet: A software application (like MS Excel) that organizes data in a grid of rows and columns, enabling storage, calculation, analysis, and retrieval. Spreadsheets are primarily used to work with numbers and text. In Excel, a spreadsheet is often called a worksheet.
* Reading data into Excel: This refers to importing data into an Excel workbook.
  + Method: You can type data directly into cells, copy-paste data from another source, or import data from external files such as Comma Separated Values (.csv) or text files (.txt).
  + Example: To import data from a .csv file, go to the Data tab in Excel, select Get & Transform Data > From Text/CSV, locate the file, and choose the appropriate delimiter (e.g., comma or semicolon).
* Basic data manipulation in Excel: This involves transforming, cleaning, reorganizing, and restructuring raw data to make it more organized and easier to interpret.
  + Example: You might sort data alphabetically or numerically, filter to show only specific entries, remove duplicate rows, or use functions to combine or split data.
* Basic functions in Excel: Excel offers predefined formulas called functions that perform specific calculations or tasks.
  + Example:
    - =SUM(A1:A10) calculates the total of numbers in the range A1 through A10.
    - =AVERAGE(B1:B20) determines the mean of values in the range B1 through B20.
    - =COUNT(C:C) counts the number of cells in column C that contain numerical values.

Spreadsheet functions to organize data

* IF function: A logical function that checks whether a condition is met and returns one value if TRUE and another value if FALSE.
  + Syntax: =IF(logical\_test, value\_if\_true, value\_if\_false).
  + Example: =IF(C2>500,"High Sales","Low Sales") checks if the value in cell C2 is greater than 500. If it is, the cell displays "High Sales"; otherwise, it displays "Low Sales".
* Nested IF: Using multiple IF functions within a single formula to test multiple conditions.
  + Example: To assign grades based on scores:  
    =IF(D2>89,"A",IF(D2>79,"B",IF(D2>69,"C",IF(D2>59,"D","F"))))  
    This formula first checks if D2 is greater than 89 for an "A", then if it's greater than 79 for a "B", and so on.
* VLOOKUP function: Searches for a value in the first column of a data range (or table array) and returns a corresponding value from a specified column in the same row.
  + Syntax: =VLOOKUP(lookup\_value, table\_array, col\_index\_num, [range\_lookup])
  + Example: =VLOOKUP(101, A2:C10, 3, FALSE) finds "101" in column A (Employee ID) and returns the value from the 3rd column (Department) in the same row.
* HLOOKUP function: Similar to VLOOKUP, but it searches horizontally in the first row of a data range and returns a value from a specified row below the lookup row.
  + Syntax: =HLOOKUP(lookup\_value, table\_array, row\_index\_num, [range\_lookup]).
  + Example: =HLOOKUP("Feb", B1:D4, 2, FALSE) finds "Feb" in the first row and returns the value from the 2nd row (e.g., Sales amount) in the same column.

Introduction to filtering, pivot tables, and charts

* Data filtering in Excel: A tool that allows you to display only the data that meets specific criteria by temporarily hiding rows that do not match.
  + Use: Filters help focus on relevant information and analyze specific subsets of data within a larger dataset.
  + Example: In a sales dataset, you can filter to show only sales for "Product X" in the "North" region. You can also filter by number conditions (e.g., sales greater than 500) or text conditions (e.g., items containing "Hoodie").
* Use of pivot tables in Excel: A powerful tool for summarizing, calculating, and analyzing large datasets dynamically.
  + Use: Identifying trends, comparisons, and patterns in data by rearranging and aggregating fields, typically including rows, columns, values, and filters.
  + Example: Analyzing monthly sales data to identify which product categories sold best in which quarter.
* Introduction to charts in Excel: Visual representations of data that help in understanding trends, patterns, and comparisons more easily.
  + Types: Common chart types include bar charts, line charts, pie charts, and scatter plots.
  + Use: Presenting data in an attractive and understandable way for reports, presentations, or quick analysis.
  + Example: A bar chart to compare sales totals across different product categories or a line chart to visualize sales trends over time.