

Pivot Tables

What are Pivot Tables used for?

- Pivot tables enable efficient data analysis, allowing users to visualize numerical data, summarize details, and identify trends quickly. With Excel and pivot tables, you can:
 - Compare product sales totals
 - Combine and remove duplicates
 - Filter out irrelevant data
 - Calculate totals or percentages

How to Make a Pivot Table in Excel

- Select your Data
 - Go to the Insert Tab>> Select Pivot Table
 - Drag and drop fields into the Rows, Columns, Values, and Filters areas within the PivotTable Fields pane.
 - Adjust and Format
-

Functions

Max/Min Functions

- **MAX(range)**: Returns the largest number in a set of values.
 - **Example:** **=MAX(A1:A10)** returns the largest value in the range A1 to A10.
- **MIN(range)**: Returns the smallest number in a set of values.
 - **Example:** **=MIN(A1:A10)** returns the smallest value in the range A1 to A10.

IF/IFS Functions

- **IF(logical_test, value_if_true, value_if_false)**: Executes a conditional test and returns one value if true and another if false.
 - **Example:** **=IF(A1>10, "Yes", "No")** returns "Yes" if A1 > 10, otherwise "No."
- **IFS(logical_test1, value_if_true1, [logical_test2, value_if_true2], ...)**: Allows multiple conditions to be evaluated in order.
 - **Example:** **=IFS(A1>10, "Large", A1>5, "Medium", TRUE, "Small")** checks conditions in sequence.

LEN Function

- **LEN(text)**: Returns the number of characters in a text string (including spaces).
 - **Example:** =LEN("Hello") returns 5.

LEFT/RIGHT Functions

- **LEFT(text, num_chars)**: Extracts a specified number of characters from the start (left) of a text string.
 - **Example:** =LEFT("Excel", 2) returns "Ex."
- **RIGHT(text, num_chars)**: Extracts a specified number of characters from the end (right) of a text string.
 - **Example:** =RIGHT("Excel", 2) returns "el."

TEXT Function

- **TEXT(value, format_text)**: Converts a number or date into text in a specified format.
 - **Example:** =TEXT(TODAY(), "MM/DD/YYYY") converts the current date to the format "12/05/2024."

TRIM Function

- **TRIM(text)**: Removes all extra spaces from a text string, leaving only single spaces between words.
 - **Example:** =TRIM(" Hello World ") returns "Hello World."

SUBSTITUTE Function

- **SUBSTITUTE(text, old_text, new_text, [instance_num])**: Replaces occurrences of text within a string.
 - **No Instances:** If the **old_text** doesn't appear, the function returns the original text.
 - **1 Instance:** By default, replaces all instances. If **instance_num** is specified, it replaces only that instance.
 - **Example:** =SUBSTITUTE("apple, apple", "apple", "orange", 2) replaces only the second "apple."

SUM/SUMIF/SUMIFS Functions

- **SUM(range)**: Adds all the numbers in a range.

- **Example:** `=SUM(A1:A10)` adds the values in A1 to A10.
- **SUMIF(range, criteria, [sum_range]):** Adds the values in a range that meet a single condition.
 - **Example:** `=SUMIF(A1:A10, ">10")` adds values greater than 10.
- **SUMIFS(sum_range, criteria_range1, criteria1, ...):** Adds values that meet multiple criteria.
 - **Example:** `=SUMIFS(B1:B10, A1:A10, ">10", C1:C10, "<20")`.

COUNT/COUNTIF/COUNTIFS Functions

- **COUNT(range):** Counts the number of numeric values in a range.
 - **Example:** `=COUNT(A1:A10)` counts numeric entries in A1 to A10.
- **COUNTIF(range, criteria):** Counts the number of cells in a range that meet a single condition.
 - **Example:** `=COUNTIF(A1:A10, ">10")` counts values greater than 10.
- **COUNTIFS(criteria_range1, criteria1, ...):** Counts cells that meet multiple conditions.
 - **Example:** `=COUNTIFS(A1:A10, ">10", B1:B10, "<20")`.

CONCATENATE Function

- **CONCATENATE(text1, [text2, ...]):** Combines multiple text strings into one.
 - **Example:** `=CONCATENATE("Hello", " ", "World")` returns "Hello World."
 - Note: The newer **CONCAT** function replaces **CONCATENATE**.

DAYS/NETWORKDAYS Functions

- **DAYS(end_date, start_date):** Returns the number of days between two dates.
 - **Example:** `=DAYS("12/31/2024", "12/01/2024")` returns 30.
- **NETWORKDAYS(start_date, end_date, [holidays]):** Returns the number of working days (excludes weekends and optional holidays).
 - **Example:** `=NETWORKDAYS("12/01/2024", "12/31/2024")`.

XLOOKUP Function

- **Purpose:** Searches a range or array for a value and returns the corresponding value from another range.
- **Syntax:**
`XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])`
 - **lookup_value:** The value to search for.
 - **lookup_array:** The range to search in.
 - **return_array:** The range to return corresponding values from.
 - **[if_not_found]:** Optional value to return if no match is found.
 - **[match_mode]:** 0 for exact match (default), -1 for exact or next smaller, 1 for exact or next larger, 2 for wildcard match.
 - **[search_mode]:** 1 for search first-to-last (default), -1 for last-to-first, 2 for binary search ascending, -2 for binary search descending.
- **Example:**
`=XLOOKUP(105, A2:A10, B2:B10, "Not Found")` searches for **105** in range **A2:A10** and returns the corresponding value from **B2:B10**, or "Not Found" if no match exists.

XLOOKUP Multiple Rows

- **Purpose:** Returns multiple values from the same or adjacent rows when matching a lookup value.
- **Syntax:** Same as standard XLOOKUP, but `return_array` can be multiple columns.
- **Example:**
`=XLOOKUP(105, A2:A10, B2:D10)` searches for **105** in **A2:A10** and returns values from **B2:D10** (multiple columns).

XLOOKUP Exact Match

- **Purpose:** Finds a value only if an exact match exists.
- **Syntax:**
Set `match_mode` to **0** (default) for exact matches.
 - **Example:**
`=XLOOKUP("John", A2:A10, B2:B10)` searches for "John" and returns the exact match from column B. Returns an error if no match is found unless `[if_not_found]` is specified.

XLOOKUP Search Order

- **Purpose:** Controls whether the function searches from top to bottom, bottom to top, or uses binary search.
- **Syntax:** Use the optional `[search_mode]` argument:
 - `1`: Search from top to bottom (default).
 - `-1`: Search from bottom to top.
 - `2` or `-2`: Binary searches for sorted data.
- **Example:**
`=XLOOKUP(105, A2:A10, B2:B10, "Not Found", 0, -1)` searches for `105` starting from the bottom of the range.

XLOOKUP Horizontal

- **Purpose:** Searches for a value in a row (instead of a column).
- **Syntax:** Same as standard XLOOKUP, but `lookup_array` and `return_array` are rows.
- **Example:**
`=XLOOKUP("Q2", B1:F1, B2:F2)` searches for "Q2" in row `B1:F1` and returns the corresponding value from row `B2:F2`.

XLOOKUP with Sum

- **Purpose:** Uses XLOOKUP to find a range of values to sum.
- **Syntax:** Use `SUM` with XLOOKUP to specify a range.
- **Example:**
`=SUM(XLOOKUP("Start", A2:A10, B2:B10):XLOOKUP("End", A2:A10, B2:B10))` sums values between the rows where "Start" and "End" appear.

VLOOKUP Function

- **Purpose:** Searches for a value in the first column of a range and returns a value in the same row from another column.
- **Syntax:**
`VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])`
 - **lookup_value:** The value to search for.
 - **table_array:** The table to search in.
 - **col_index_num:** The column number to return a value from.
 - **[range_lookup]:** TRUE for approximate match, FALSE for exact match.

- **Example:**

`=VLOOKUP(105, A2:D10, 3, FALSE)` searches for **105** in the first column of **A2:D10** and returns the value from the third column.

Key Differences Between XLOOKUP and VLOOKUP

1. **Search Direction:** XLOOKUP can search both vertically and horizontally; VLOOKUP only searches vertically.
2. **Flexibility:** XLOOKUP allows you to define separate lookup and return arrays; VLOOKUP requires the return column to be in the same range.
3. **Default Match:** XLOOKUP defaults to an exact match; VLOOKUP defaults to an approximate match unless specified.
4. **Performance:** XLOOKUP supports advanced search modes, such as searching backward or using binary search for sorted data.

Wildcard Explanation

A **wildcard** is a special character used in search operations to represent one or more unknown characters. Wildcards are commonly used in Excel, databases, and other software applications to perform flexible matching, especially when the exact value is partially known or variable.

Common Wildcards

1. **Asterisk (*):**

- Represents **any number of characters**, including none.
- **Example:**
 - Searching for `"Jo*"` matches `"Jo"`, `"John"`, `"Joseph"`, and `"Journal"`.
 - Searching for `"*ing"` matches `"Running"`, `"Swimming"`, and `"Eating"`.

2. **Question Mark (?):**

- Represents **exactly one character**.
- **Example:**
 - Searching for `"J?hn"` matches `"John"` and `"Jahn"` but not `"Johann"`.
 - Searching for `"T?m"` matches `"Tom"`, `"Tim"`, but not `"Team"`.

Use Cases for Wildcards

1. Flexible Lookups:

- Used in functions like **XLOOKUP**, **COUNTIF**, and **SEARCH** to find values that partially match a specified pattern.

2. Data Filtering:

- Applied in filters to quickly locate rows with matching patterns.

3. Text Matching:

- Helpful in finding and replacing data when the exact value isn't known (e.g., replacing **"*Smith"** with **"Dr. Smith"**).

Wildcard Behavior in Excel

- **Functions Supporting Wildcards:**

XLOOKUP, **COUNTIF**, **COUNTIFS**, **SEARCH**, **MATCH**, and **SUMIF** allow wildcards.

- Example: **=COUNTIF(A1:A10, "*apple*")** counts cells containing "apple" anywhere in the text.

- **Escaping Wildcards:**

If you need to search for a literal ***** or **?**, use a tilde (**~**) before the wildcard.

- Example: **=COUNTIF(A1:A10, "data~*")** searches for the text **"data*"**.
-

Conditional Formatting

Conditional Formatting Overview

- **Purpose:** Conditional formatting helps highlight trends, patterns, or outliers in your data, making it easier to analyze and interpret large datasets. It's widely used to visualize performance, identify key values, and highlight anomalies.
- **Flexibility:** You can use built-in rules or create custom rules to suit specific analytical needs, making it a powerful tool for data visualization and analysis.

Icon Sets

- **Use Cases:**
 - **Performance Tracking:** Display arrows (up, down) to track performance over time.

- **Risk Assessment:** Use traffic light icons (red, yellow, green) to represent varying levels of risk.
- **Customization:** Adjust thresholds and values for each icon to represent the correct data range. This can help visualize trends in performance or status.

Color Scales

- **Best for:** Visualizing distributions of data, like sales or temperatures.
- **Gradient Variations:**
 - **Two-color scales:** Effective when comparing two extremes (e.g., high vs. low).
 - **Three-color scales:** Ideal for datasets with a natural middle value (e.g., normal, above average, below average).
- **Advanced Use:** Combine with data trends to highlight variance or outliers, such as cells deviating significantly from the mean.

Gradient Fills

- **Purpose:** Applies a gradient of colors within a single cell. Often used to emphasize data trends within a cell or highlight variations in values.

Top/Bottom Rules

- **Best for:**
 - **Identifying outliers:** Quickly highlight the top 10 or bottom 10 items (or percentages) to find the highest or lowest values in your dataset.
 - **Summary Insights:** Highlight top-performing or worst-performing categories (e.g., highest sales, lowest returns).
- **Performance Metrics:** Combining Top/Bottom rules with icon sets enhances data visualization, allowing key performers to stand out.

Highlight Cells Rules

- **Useful For:**
 - **Spotting Duplicates:** Helps in data cleaning by identifying repeated entries.
 - **Highlighting Specific Text:** Easily locate cells containing specific text or keywords (e.g., status, category).
- **Other Rules:**
 - **Greater Than/Less Than:** Highlights values above or below a certain threshold (e.g., sales targets, profit margins).
 - **Between:** Identifies values within a specific range, useful for grouping data.

- **Equal To:** Highlights exact matches, useful for categorical data or specific criteria.

Custom Rules

- **Advanced Logic:** Leverage Excel formulas for complex conditions. For example, highlight data points based on multiple criteria (e.g., sales greater than \$1000 but only for Region A).
- **Formula Examples:**
 - `=AND(A1>1000, B1="Region A")` highlights sales greater than 1000 for Region A.
 - `=OR(A1="Yes", B1="No")` highlights rows where either condition is met.
- **Remove \$ (relative vs. absolute references):** Removing the dollar sign allows the rule to apply dynamically to each cell in the range (relative reference), rather than applying a fixed rule to each cell (absolute reference).

How to Delete Rules

- **Manage Multiple Rules:** Data analysts often work with complex datasets that require multiple formatting rules. Use **Manage Rules** to view, delete, or edit any existing rules in the current worksheet or workbook.
- **Clear Rules for a Range:** Quickly remove all formatting from a selected range if the formatting becomes too complex.
- **Clear Rules from Entire Worksheet:** Clears all conditional formatting from the entire worksheet for a fresh start.

Additional Tips

- **Use Conditional Formatting Sparingly:** Too many rules can overwhelm the data and make analysis difficult. Focus on the most important insights.
- **Visual Consistency:** Be consistent with color choices and icon styles. Make sure colors are meaningful (e.g., red for negative, green for positive), and icons should be intuitive and easy to interpret.
- **Contextual Use:** Combine conditional formatting with other data visualization tools like charts or PivotTables to provide better context to your analysis.

Charts

- Charts visually represent data, making it easier to identify patterns, trends, and insights.

- To create a chart:
 1. Select the data range.
 2. Go to the **Insert** tab and choose a chart type (e.g., column, line, pie).
- Charts can be customized with titles, labels, legends, and axes.

Switch Row/Column

- **Switch Row/Column** swaps the data displayed on the X-axis (horizontal) and Y-axis (vertical).
- Useful when the default chart arrangement does not match the desired visualization.
- Steps:
 1. Select the chart.
 2. Go to the **Chart Design** tab.
 3. Click **Switch Row/Column** in the **Data** group.

Change Chart Type

- Allows you to change the style of the chart (e.g., from a column chart to a line chart).
- Steps:
 1. Select the chart.
 2. Go to the **Chart Design** tab.
 3. Click **Change Chart Type** and select the desired style.
- Use this feature to better match your data with the most appropriate chart.

Filter

- Filters in charts let you display specific data points while hiding others.
- Steps:
 1. Click on the chart.
 2. Use the **Filter** button (funnel icon) in the chart's corner.
 3. Uncheck or check data series or categories to adjust the display.
- Helps focus on relevant data for analysis.

Chart Styles -> Color

- **Chart Styles** provide pre-designed layouts and color combinations for charts.
- Steps:
 1. Select the chart.
 2. Go to the **Chart Design** tab.
 3. Click **Chart Styles** to choose a layout or color palette.
 4. To change colors, click **Change Colors** and pick from the available themes.
- Enhances chart aesthetics and ensures consistency with presentation themes.

Data Cleaning

Data cleaning is essential for effective data analysis. Excel is a powerful tool for performing various data cleaning tasks.

Common Data Cleaning Techniques

- 1. Removing Duplicates:**
Duplicate entries should be eliminated to maintain data integrity.
- 2. Formatting and Standardization:**
Ensure consistent text formatting (e.g., capitalization).
- 3. Dealing with Extra Spaces:**
Trim unnecessary spaces from cells to clean up the data.
- 4. Handling Currencies:**
Convert currency formats to numeric values for calculations.
- 5. Date Formatting:**
Ensure that all date entries are uniform in format.

Practical Steps During Data Cleaning

- Use Excel's "Remove Duplicates" feature.
- Insert new columns for corrections without altering original data.
- Utilize functions like `UPPER()`, `LOWER()`, and `TRIM()` to standardize text entries.
- Identify and fix common data entry errors (e.g., typos in categorical variables).
- Delete unnecessary columns to focus on relevant data.

Backup and Documentation

- Always keep a backup of original data for reference.
- Use working documents for data cleaning to avoid losing raw data.