Conditional Formatting

WHEN IS THIS USEFUL?

Conditional formatting can be useful to help identify data that needs attention, such as duplicate data, numbers above or below a target, or cells containing a specific value.

HOW DOES THIS WORK?

Highlight the column you wish to apply conditional formatting to. Navigate to the Home menu and select "Conditional Formatting" in the Styles section. Choose one of the pre-set option or New Rule. Follow the prompts the establish your rule and formatting.

TAKE IT FURTHER

Conditional Formatting isn't limited to font and fill color. You can use Icons in addition to or instead of the cell values to create easy visual guides for yourself or other end users.

Data Validation

WHEN IS THIS USEFUL?

Excel's data validation features can be useful in two scenarios. This first is when you are working with an existing data set and want to be able to quickly identify data that was entered incorrectly, such as zip codes that aren't 5 digits long or email addresses that were entered into the phone number column.

The second scenario where data validation can be a major game changer is when you are setting up a spreadsheet for data entry and you want to limit the responses or type of responses that will be accepted in each column.

HOW DOES THIS WORK?

Highlight the column or cells you want to apply the validation to. Then in the Data Menu, select "Data Validation" in the Data Tools section and choose the data validation option. Follow the prompts to set the parameters of acceptable entries or link to a list of pre-set options.

TAKE IT FURTHER

Once your validation parameters are set, you can quickly flag, and therefore fix, bad data by selecting "Circle Invalid Data" from the Data Validation drop-down menu. To remove the circles, either fix the invalid data or select "Clear Validation Circles".

If Error Function

WHEN IS THIS USEFUL?

Sometimes errors pop up, especially if you have missing data in your dataset. If you are trying to calculate a percent between two cells and you have missing data for the denominator, it will return an error. You can use the IFERROR function to handle errors in a formula. IFERROR returns a value you specify if a formula evaluates to an error; otherwise, it returns the result of the formula.

HOW DOES THIS WORK?

To use the IFERROR formula you essentially wrap your existing formula in the IFERROR statement so it looks like this: =IFERROR(desired_formula, value_if_errpr).

Example: =IFERROR(B3/C3,"") where B3= numerator and C3=denominator. If C3 is blank, it errors the division and since you have used the if error function, it will return a blank because we put "" in the value if error part of the function. If C3 was not blank and thus did not error out the division formula, then it would return the calculated number.

If Then Functions

WHEN IS THIS USEFUL?

If statements can be useful when you need to recode existing data into something more useful for analysis or visualization, such as responses to a Likert scale survey question that are displaying as a numeric value, but you want the column to display the corresponding strength of agreement.

HOW DOES THIS WORK?

The basic IF formula is =IF(logical_test, value_if_true, value_if_false). So continuing with the example above, we might use the statement =IF(A2=1, "Strongly Disagree", "Agree). If the respondent is showing a 1 as their answer to the survey question, then they strongly disagreed with that statement.

TAKE IT FURTHER

There may be times where you need more than one condition to be true or one of multiple possible conditions to be true. In that case try an If-And statement: =IF(AND(logical_test1, logical_test2,...), value_if_true, value_if_false) or If-Or statement: IF(OR(logicEal_test_1, logical_test_2,...), value_if_true, value_if_false).

Joining Data from Multiple Cells

WHEN IS THIS USEFUL?

There may be times when you need to pull data from multiple columns into a single cell, such as combining first and last name into a full name column or street, city, state, and zip into a full address column.

HOW DOES THIS WORK?

The simplest way to do this is to use the ampersand (&) and quotation marks (") along with your reference cells. Simply start with an equal sign (=), select the first reference cell followed by &". Then add any spacing or characters that should come between the reference cells. Next enter "& before selecting your second reference cell. Adding a third or forth piece of information? Simply sandwich reference cells that come in the middle of your equation between "& and &".

A formula joining two columns of information may look like this: =reference1&" "&reference2
A formula joining three columns of information may look like this: =reference1&" "&reference2&" "&reference3

TAKE IT FURTHER

Another option for joining information from multiple cells is =CONCAT(), which will pull the information together without spaces or formatting. Or, if you have a lot of these transformations to do, you can use PowerQuery to Add Custom columns.

Nested If Then Functions

WHEN IS THIS USEFUL?

Nested If statements can be a powerful function when there is not necessarily a clear response if the logical test of your first if statement is false, such as the Likert scale example above. Often these scales are a range of 1 to 4 or 5, with each number corresponding to a certain amount of agreement so if a response isn't equal to 1, there are multiple other ways the respondent could have felt about the statement.

HOW DOES THIS WORK?

The base of a nested if statement is simple an if statement when the value_if_false is a second if statement.

Example: Continuing with the Likert scale example and assuming a five-point scale, a nest if statement converting numeric values to level of agreement might look like this =if(A2=1, “Strongly Disagree”, IF(A2=2, “Disagree”, IF(A2=3, “Neutral”, IF(A4=4, “Agree”, “Strongly Agree”))))

TAKE IT FURTHER

Is your nested if statement getting long and unwieldly, try using PowerQuery to create a conditional column or group/bin responses.

Parsing Date

WHEN IS THIS USEFUL?

There may be times when too much information is contained within a single column and you want to isolate chunks of that information into their own columns to aid your analysis, such as when a full address is contained with a single cell but you want to see distribution across cities or when zip codes are in zip+4 format but you really just need to use the first five digits.

HOW DOES THIS WORK?

Highlight the column you wish to split. Navigate to the Data menu and select “Text to Columns” in the in the Data Tools section. Choose whether you want to split your column at the location of a specific character (Delimited) or after a specific number of characters (Fixed width). Then follow the prompts to select that specific character or number of characters.

TAKE IT FURTHER

Parsing data can be very helpful when cleaning data. Try delimiting at the @ for a copied column of email address to verify there aren’t any typos in the domains potentially limiting the number of bad email addresses in your list. Or try isolating the first seven characters of a copy column of street address to help identify potential duplicates while ignoring possible variations of street, drive, avenue, etc.

Pivot Tables

WHEN IS THIS USEFUL?

A Pivot Table is used to summarize, sort, reorganize, and group data by allowing you to transform columns into rows and rows into columns.. You can quickly count, total, or average your data across

multiple categories within the same column or multiple columns without needing complicated formulas.

HOW DOES THIS WORK?

First, ensure your data is formatted with a single row of unique headers. Then, select the first header or highlight the full range of data you wish to use in your pivot table. In the Insert menu, select "Pivot Tables" in the Tables section and choose "from Table/Range" from the drop down menu. You will be prompted to choose where you want your pivot table to be place, a new worksheet or a specific cell in an existing worksheet.

Once you have you have your location set, a new menu will open on the right-hand side. In this menu, you can drag your original column headers Filters, Columns, Rows, and Values boxes in the bottom of the menu to create your pivot table. If the values aren't displaying correctly, open the drop down menu for that value and select "Value Field Settings". You can choose a number of summarization functions, such as count, sum, and average.

TAKE IT FURTHER

It is best practice to format your source data as an active table (Home>Style>Format as Table). This makes your pivot table dynamic and responsive. It will automatically pull in new rows of information as they are added and will refresh without prompting by you.

PowerQuery

WHEN IS THIS USEFUL?

Power Query is a data transformation and data preparation engine build directly into Excel. If you need to append or merge data from multiple worksheets or workbooks or if you need to make a number of transformations to your data if may be a better choice than implementing multiple pivot tables and complicated formulas, as Power Query records every action taken so you can quickly go back to correct errors or make updates without having to redo multiple steps.

HOW DOES THIS WORK?

If you are using data contained within the same workbook, select the first header cell of your range or table. In the Data Menu, select From "Table/Range" in the Get & Transform Data section. Pulling in data from outside the workbook? Select "Get Data", and then select your source. Make your transformations within the Power Query window and choose "Close and Load" to bring your new

query into a worksheet.

TAKE IT FURTHER

Power Query can do so much, it is difficult to cover everything here. If you are interested in building this skill, we recommend checking out the Power Query content on Microsoft Learn (https://learn.microsoft.com/en-us/power-query/) or a Power Query training on LinkedIn Learning, such as Excel: Power Query (Get & Transform) or Excel: Power Query for Beginners.

Removing Duplicates

WHEN IS THIS USEFUL?

Typically when working with data about people or events, we want to make sure each is accounted for only once, so an important first step when cleaning data is to remove duplicate rows of data.

HOW DOES THIS WORK?

If you have a unique identifier, highlight your full range of data or table, navigate to the Data Menu and select "Remove Duplicates" from the Data Tools section. Select the column with your unique identifier and hit ok. This will remove all duplicate rows, keeping just the 1st instance.

TAKE IT FURTHER

If your data does not have a unique identifier for each row, you may need to use a more manual method of identifying and removing duplicates. In this case, highlight the column you believe might have duplicate information. Then in the Home Menu, select "Conditional Formatting" in the Styles section, and choose "Highlight Duplicates". You can then filter that column on fill color and sort A to Z to review and verify if the duplicates are truly repeated data or unique.

Summarization Functions

WHEN IS THIS USEFUL?

Summarization formulas allow you to carry quick calculations that give you a big-picture view of the information within your dataset. You can use these formulas to sum, count, find central tendency, or find minimum and maximum to calculate range.

HOW DOES THIS WORK

The following is a list of formulas for commonly used summarization functions.

Function	Formula	Description

AVERAGE	=AVERAGE(number1, number2, ...)	Returns the average (arithmetic mean) of its arguments, which can be numbers or names, arrays, or references that contain numbers.
AVERAGEIF	=AVERAGEIF(range, criteria, average_range)	Finds the average (arithmetic mean) for cells specified by a given condition or criteria.
COUNT	=COUNT(value1, value2, ...)	Counts the number of cells in a range that contain numbers
COUNTA	=COUNTA(value1, value2, ...)	Counts the number of cells in a range that are not empty.
COUNTBLANK	=COUNTBLANK(value1, value2, ...)	Counts the number of empty cells in a specified range of cells.
COUNTIF	=COUNTIF(range, criteria)	Counts the number of cells withing a range that meet the given condition.
MAX	=MAX(number1, number2, ...)	Returns the largest value in a set of values. Ignores logical values and text.
MAXIFS	=MAXIFS(Max_range, criteria_range, criteria, ...)	Returns the maximum value among cells specified by a given set of conditions or criteria.
MIN	=MIN(number1, number2, ...)	Returns the smallest number in a set of values. Ignores logical values and text.

MINIFS	=MINIFS(min_range, criteria_range, criteria,...)	Returns the minimum value among cells specified by a given set of conditions or criteria.
SUM	=SUM(number1,number2,...)	Adds all the numbers in a range of cells.
SUMIF	=SUMIF(range, criteria, sum_range)	Adds the cells specified by a given condition or criteria.

VLOOKUP Function

WHEN IS THIS USEFUL?

You may find that you have connected data spread across multiple worksheets. Use the VLOOKUP (Vertical Look-up) function to pull data from one worksheet into another worksheet based on connected unique identifiers.

HOW DOES THIS WORK?

To use the VLOOKUP function, you need a few pieces of information: the unique value that ties your row of information to data in a different worksheet (the look-up value), a range of cells containing both the look-up value and the data you wish to bring in to your row, and the column number of that data, where column 1 is the column your look-up value will be matched to.

=VLOOKUP(lookup_value, table_array, column_index_num, [range lookup])

TAKE IT FURTHER

You are not just limited to matching look-up values within a column. Use the HLOOKUP function to match your look-up value within a row or the XLOOKUP function to match your look-up value across columns and rows.