Data Analysis in Excel

Recommendations and Sample Code

*Sign up for free:* [*https://signup.live.com/?lic=1*](https://signup.live.com/?lic=1)

# Tips and Tricks

## Use macros!

Macros can be used to auto-populate sheets and auto-copy sheets. It’s also possible to write custom functions.

Documentation of VBA is good and stack exchange and other related forums have lots of questions/answers to help debug and find solutions.

If you find a particular command isn’t working for what you need, search around for another option.

**Note:** If there is a LOT of data, do not run from VBA editor (this may cause Excel to crash). Instead, run from the Macros menu (see photo to the right)

## Use tables

Once you’ve given the dimensions of the table (headers and first column), equations auto-populate and auto-update.

As you perform calculations with tables, you can use tab-complete for headers to make the process much faster.

Tables also allow you to reference data, either an individual cell, entire columns, or multiple columns, by header names. If the size of the table changes, or you add/remove columns, equations still function as expected.

## Case-insensitive and partial string comparisons

To do case-insensitive string searches, or to do partial string comparisons, use "\*thing\*" to search for words that contain "thing".

For example, if you’re looking for responses that reference “accessibility” and similar, you can use: “\*access\*”, which will look for any string containing “access”, including accessibility, accessible, and access.

## Utilize horizontal capability of sheets.

Avoid doing calculations at the bottom of tables so that you can allow for more data to come in without disrupting or deleting your prior work. Instead, work from left-to-right and add analysis tables to the right of your raw data tables.

In this way, it is quick and easy to update tables as new data comes in, hooray! See the “Copy()” macros function for more details.

# Helpful Commands

## Excel Functions

The following are common functions used in data analysis. More information on functions, including examples, in Excel can be found using the search feature inside Excel.

1. **Basic statistics**

The following are formulas for conducting basic statistical analysis. Note that formulas are not case sensitive, e.g. “AVERAGE” gives the same as “average” and “Average”.

* Count
  + Syntax: COUNT(value1, [value2], ...)
  + Description: The COUNT function counts the number of cells that contain numbers, and counts numbers within the list of arguments. Use the COUNT function to get the number of entries in a number field that is in a range or array of numbers.
* Average
  + Syntax: AVERAGE(number1, [number2], ...)
  + Description: Returns the average (arithmetic mean) of the arguments. For example, if the range A1:A20 contains numbers, the formula =AVERAGE(A1:A20) returns the average of those numbers.
* Median
  + Syntax: MEDIAN(number1, [number2], ...)
  + Description: Returns the median of the given numbers. The median is the number in the middle of a set of numbers.
* Mode
  + Syntax: MODE(number1,[number2],...)
  + Description: MODE returns the most frequently occurring, or repetitive, value in an array or range of data.
* Minimum
  + Syntax: MIN(number1, [number2], ...)
  + Description: Returns the smallest number in a set of values.
* Maximum
  + Syntax: MAX(number1, [number2], ...)
  + Description: Returns the largest value in a set of values.
* Standard deviation
  + Syntax: STDEV.S(number1,[number2],...)
  + Description: Estimates standard deviation based on a sample (ignores logical values and text in the sample).
  + The standard deviation is a measure of how widely values are dispersed from the average value (the mean).

1. **Countif(…)**

CountIf allows you to count up data in cells that meet certain criteria. You can add multiple countif() statements together to count multiple criteria. You can also subtract if criteria overlap (e.g. if looking for "ca" as a state but want to exclude canada)

**Examples**:

* + 1. 'Count a single response if it includes "CA" or "ca" '

=COUNTIF(Table8[@[What state do you live in?]], "\*CA\*")

1. 'Count a single response If it includes California or some variation, including CA, ca, cali, etc. '

=COUNTIF(Table8[@[What state do you live in?]], "\*CA\*") + COUNTIF(Table8[@[What state do you live in?]], "\*cali\*")

1. 'Count a single response if it includes California or some variation, but discount a response if it includes Canada'

=COUNTIF(Table8[@[What state do you live in?]], "\*CA\*") + COUNTIF(Table8[@[What state do you live in?]], "\*cali\*") **-** COUNTIF(Table8[@[What state do you live in?]], "\*canada\*")

1. 'Count all responses if they include "CA"'

=COUNTIF(Table8[What state do you live in?], "\*CA\*")

1. **Logic Statements**

The following are basic logical statements to separate and analyze data based on conditionals. Many of the logical functions can be used together.

* If
  + Syntax: IF(logical\_test, [value\_if\_true], [value\_if\_false])
    - Example: =IF(C2=”Yes”,1,2)
      * Result: IF(C2 = Yes, then return a 1, otherwise return a 2).
  + Description: IF allows you to make logical comparisons between a value and what you expect. An IF statement can have two results. The first result is if your comparison is True, the second if your comparison is False.
* And
  + Syntax: AND(logical\_1, [logical\_2],…)
  + Description: Returns TRUE if **all** the arguments are true.
* If error
  + Syntax: IFERROR(value, value\_if\_error)
  + Description: Returns a value you specify if a formula evaluates to an error; otherwise, returns the result of the formula
* Not
  + Syntax: NOT(logical)
  + Description: Reverses the logic of its argument.
* Or
  + Syntax: OR(logical\_1, [logical\_2],…)
  + Description: Returns TRUE if **any** argument is true.
* XOR
  + Syntax: XOR(logical\_1, [logical\_2],…)
  + Description: Returns a logical exclusive OR of all arguments

## Macros VBA commands

## 

1. **Add() Function**

Examples:

**'Name the new sheet'**

Sheets.Add.Name = "NameOfNewSheet"

**'Add a Single Column to the new sheet'**

Sheets("CopyFromHere").Range("A:A").Copy Destination:=("NewSheetName").Range("A:A")

**'Add Multiple Columns'**

Sheets("CopyFromHere").Range("A:C").Copy Destination:=("NewSheetName").Range("A:C")

**'Add Multiple Columns that are not next to each other'**

Sheets("CopyFromHere").Range("A:A,C:C").Copy Destination:=("NewSheetName").Range("A:A,B:B")

1. **Copy() Function**

Examples

**'This will copy all data in Column A from a sheet (CopyFromHere) into an existing sheet (CopyToHere)'**

Sheets("CopyFromHere").Range("A:A").Copy \_

Destination:=Sheets("CopyToHere").Range("A:A")

**'This will copy all data in Column F from a sheet (CopyFromHere) into an existing sheet (CopyToHere)'**

Sheets("CopyFromHere").Range("F:F").Copy \_

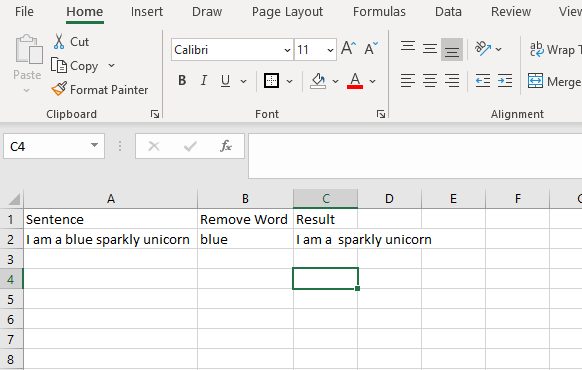
Destination:=Sheets("CopyToHere").Range("B:B")

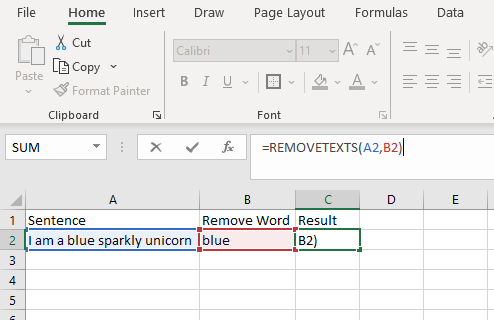
1. **Custom Function: RemoveTexts**

This function removes text contained in a selection and returns the remaining text.

For example, in the sentence: "I am a blue sparkly unicorn", if you wanted to remove "blue", you would put the sentence in one cell, the word to remove in another, then use:

=removetexts(CellWithSentence, CellWithWordBlue)



****

**Full VBA for function:**

Function REMOVETEXTS(strInput As String, rngFind As Range) As String

Dim strTemp As String

Dim strFind As String

strTemp = strInput

For Each cell In rngFind

strFind = cell.Value

strTemp = Replace(Replace(strTemp, strFind, ""), ";", "")

Next cell

REMOVETEXTS = strTemp

End Function

## Troubleshooting

1. Cannot add two of the same sheet name.
   1. If you get an error: "that name is already taken…", it means that you have already created that sheet. Go back to Excel view and check that the sheet looks as expected. If not, delete and try again.
      1. **Note**: this error message will result in a blank sheet, be sure to delete!
   2. If you are testing code, just delete the sheet and run again (sheet will populate each time the script is run)
2. Search for the problem! Lots of good Q & A online as well as example code.
   1. [**Stack Overflow**](https://stackoverflow.com/search?q=microsoft+excel&s=f9e8c948-342d-46e4-ad8e-419c68ed4c0f) is a great resource.