Lab 9

Table of contents

Excel basics

Lab class

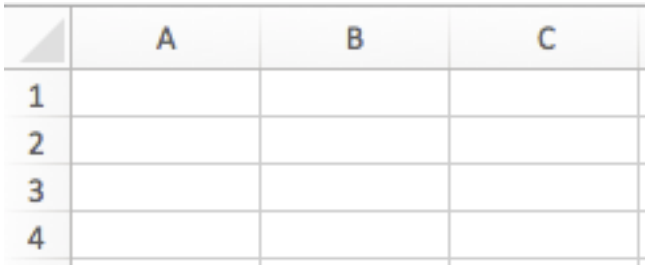
To analyse data, it will be helpful to be a confident user of Excel. In fact, no matter what career path you choose, you will probably need to use a spreadsheet application from time to time. Therefore, having good Excel skills will not only help you in your studies, but will also be a great asset for future internship or job applications.

Please note that you will not need to remember the formulas you are going to use in this lab. However, you should be able to find them in this document and adapt them to a new analysis at a later point in time.

If you have very little previous experience using Excel, you might also want to complete an introductory Excel course online. Many of these courses are available for free.

Columns, rows, and cells

At the most basic level, you need to understand the layout of worksheets in columns, rows, and cells. In Excel, columns have letters, rows have numbers:



Every cell is addressed by the combination of column letter and row number. For example, the first cell is referenced like this: `A1`.

Cut, copy, and paste

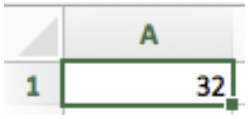
The usual shortcuts also work in Excel (`Cmd` for macOS, `Ctrl` for Windows):

- Cut: `Cmd + X` or `Ctrl + X`
- Copy: `Cmd + C` or `Ctrl + C`
- Paste: `Cmd + V` or `Ctrl + V`

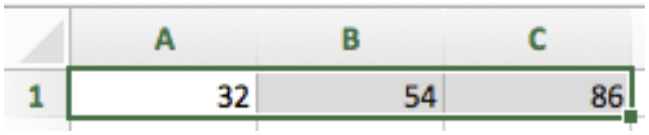
Note that there is a very useful variant of paste: "Paste special". You can read more about "Paste special" here.

Selecting content

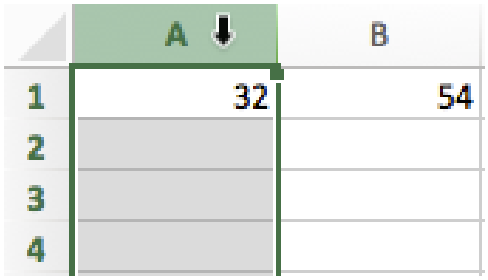
To select one cell, click on it.



To select adjacent cells, click on the first cell, hold the mouse button down, and drag the mouse to the target cell.

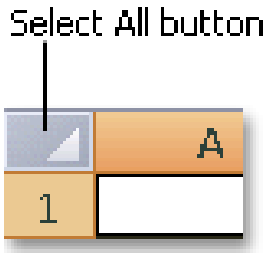


To select a column, click on the cell containing the letter (note how the mouse pointer changes into a downward-pointing arrow).



Selecting rows works in an analogous way.

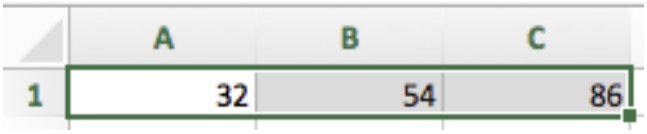
To select all cells in a worksheet, click on the `Select All` button:



You can read more about selecting cell contents here.

Cell ranges

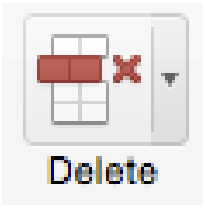
Cell ranges are indicated by colons. In this example, the selected cell range would be written as `A1:C1`:



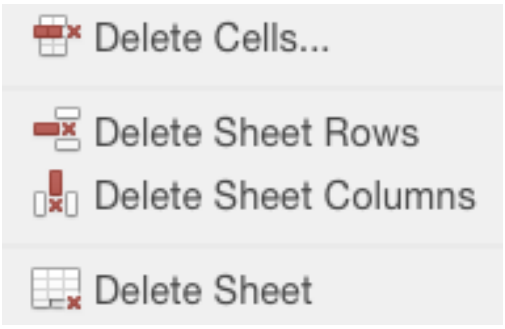
Cell ranges can span rows and columns. For example, `A1:C3` is a range of nine cells.

Deleting cells, rows or columns

Select the cells, rows or columns, then click on `Delete`:



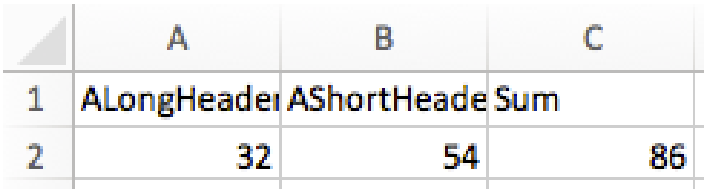
Alternatively, you can use the `Delete` dropdown menu:



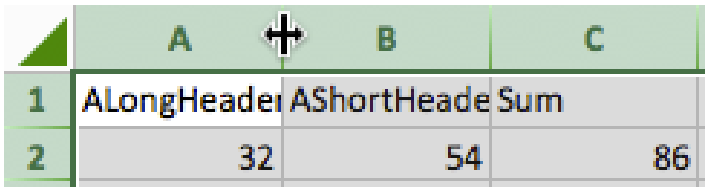
Note that `Delete` will not only delete the content of the cells/rows/columns, but the cells/rows/columns themselves.

Automatically adjusting column width

The default column width in Excel is 8.43 characters using the default font. This is somewhat annoying, as it will frequently only display part of your headers:



To quickly autofit the width of all columns on the worksheet, click the `Select All` button, and then double-click any boundary between two column headings (note how the mouse pointer turns into a double arrow):



	A	B	C
1	ALongHeade	AShortHeade	Sum
2	32	54	86

The result:

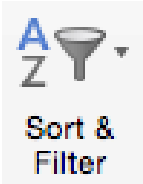
	A	B	C
1	ALongHeaderForThisColumn	AShortHeader	Sum
2	32	54	86

This is especially useful for PsychoPy output files as relevant information frequently appears at the end of the column header (e.g., .corr or .rt).

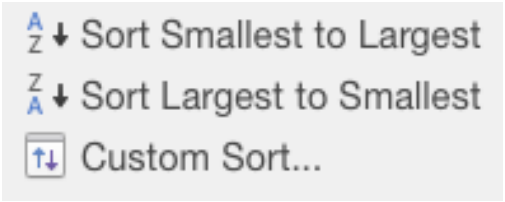
You can read more about changing column widths and row heights here on this topic.

Sorting data

Click on this icon to sort (or filter):

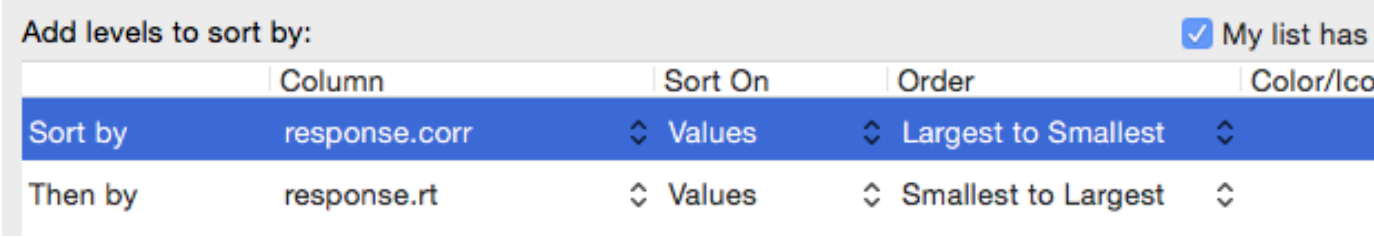


You will almost always need `Custom Sort...`:



If your worksheet has headers, make sure `My list has headers` is ticked:

Custom sort allows you to add multiple levels of sorting criteria. This example would first sort by `response.corr` (first listing all correct trials), and then by `response.rt` (first listing the fastest trials):



The result might look like this:

	A	B
1	response.rt	response.corr
2	104	1
3	634	1
4	723	1
5	4396	1
6	549	0
7	583	0

Note: You need to make sure to select **all** columns before sorting (e.g., using the `Select All` button), otherwise your columns might end up being out of sync (e.g., RTs might no longer be associated with the matching accuracy information). Excel might present you with a popup warning if you fail to select all columns.

Data formats

Data in Excel cells can have any one of a number of different formats. The most relevant for us are text and numbers (more information about number formats is available here). There is a simple way to tell text and numbers apart without explicitly checking the format: Text in a cell is left-aligned, numbers are right-aligned. Note that some functions (e.g., `COUNT`) only work on cells formatted as numbers.

Relative vs. absolute cell references

There are two types of cell references in Excel, relative and absolute. Relative references are the default. Absolute references are indicated by the `$` sign.

- `A1` is a relative reference to the cell `A1`.
- `$A$1` is an absolute reference to the cell `A1`.

Please note that you can mix relative and absolute cell references:

- `A$2` is a relative reference to column `A`, and an absolute reference to row 2.
- `$A2` is an absolute reference to column `A`, and a relative reference to row 2.

How cell references work is explained in Section .

Excel formulas and functions

Lab class

Formulas

Important basics about formulas in Excel:

- Formulas in Excel always begin with an equals sign: `=`
- The formula is always typed into the cell where the answer should appear.
- The formula is completed by pressing `Enter` on the keyboard.

Functions

Formulas can make use of predefined functions in Excel. We will only describe a small set of functions that will be fundamental to our analyses. For more information, see this page about Excel functions sorted by category or this page about Excel functions in alphabetical order.

If you want help on functions in Excel, click on the function symbol fx:

This will open the Formula Builder, providing you with information on the available functions, and if you double-click on a specific function, help with building the formula. Alternatively, go to "Help" □ "Excel Help", or search online.

IF

`IF` is one of the most important logical functions in Excel. It works in the following way:

```
IF(something is True, then do something, otherwise do something else)
```

An `IF` statement can have two possible outcomes, depending on whether the condition is True or False. A potential use case for this function could be to create a column that includes only correct RTs.

Imagine you have the following PsychoPy output:

	A	B	C
1	response.rt	response.corr	only_correct_rts
2	634	1	
3	549	0	
4	583	0	
5	723	1	

The first column contains all RTs, and the second column information about accuracy (remember that PsychoPy codes correct responses as 1, and incorrect responses as 0). Your aim is to have only RTs associated with correct responses in your third column.

You can use the following formula in cell `C2` to achieve this:

```
=IF(B2=1, A2, "")
```

This states that:

- If `B2` equals 1 (i.e., if the trial is correct)
  - Then `C2` should be equal to `A2` (basically copying the value from `A2` to `C2`)
- If this is not the case (i.e., if the trial is not correct) [this part is implicit to the function]
  - Then `C2` should remain empty (indicated by `""`)

If we apply this formula to all rows in our worksheet (by copying it to the adjacent cells - see below for details), we get the following output:

	A	B	C
1	response.rt	response.corr	only_correct_rts
2	634	1	634
3	549	0	
4	583	0	
5	723	1	723

Column `C` now only contains RTs associated with correct trials.

Note that you can create more complex statements by combining multiple conditions using `AND` or `OR`. For example, you might not only want to remove RTs from incorrect trials, but also RTs that are extremely fast (presumably representing anticipatory responses) or extremely slow (presumably representing lapses of attention), independent of their accuracy. Based on previous research, you might consider RTs below 150 ms as anticipatory, and RTs over 4000 ms as lapses of attention. You could use the following formula to achieve your aim:

```
=IF(AND(B2=1, A2>=150, A2<=4000), A2, "")
```

This states that:

- If `B2` equals 1 **and** A2 is equal to or above 150 ms **and** A2 is equal to or less than 4000 ms
  - Then `C2` should be equal to `A2` (basically copying the value from `A2` to `C2`)

- If this is not the case
  - Then `C2` should remain empty (indicated by `""`)

Result:

	A	B	C
1	response.rt	response.corr	usable_rts
2	634	1	634
3	549	0	
4	583	0	
5	723	1	723
6	104	1	
7	4396	1	

Quite frequently, you will take into account an additional column indicating the condition. For example, there might be an additional column coding the congruency of the stimulus, with the levels `con` and `incon`. If you wanted to have separate columns for your congruent and incongruent RTs, you could do this as illustrated below.

Create a column with all congruent RTs eligible for further analysis:

```
=IF(AND(B2=1, A2>=150, A2<=4000, C2="con"), A2, "")
```

Create a column with all incongruent RTs eligibel for further analysis:

```
=IF(AND(B2=1, A2>=150, A2<=4000, C2="incon"), A2, "")
```

Please note that letter strings need to be in quotes.

Result:

	A	B	C	D	E
1	response.rt	response.corr	congruency	usable_rts_con	usable_rts_incon
2	634	1	con	634	
3	549	0	incon		
4	583	0	con		
5	723	1	con	723	
6	104	1	incon		
7	4396	1	con		
8	723	1	incon		723
9	443	1	con	443	
10	632	1	incon		632
11	652	1	incon		652

SUM

`SUM` is another relevant function. You can add individual values, cell references or ranges or a mix of all three.

For example:

`=SUM(A2:A10)`: Add the values in cells `A2` to `A10`.

`=SUM(A2:A10, C2:C10)`: Add the values in cells `A2` to `A10` and in cells `C2` to `C10`.

AVERAGE

`AVERAGE` is another function you will frequently need. It returns the average (arithmetic mean). For example, if the range `A1:A20` contains numbers, the formula `=AVERAGE(A1:A20)` returns the average of those numbers.

Please note that the value `0` will be treated like any other number by this function (thus affecting your mean!), whereas empty cells are ignored.

Variants of `AVERAGE` are `AVERAGEIF` and `AVERAGEIFS`. These variants allow the specification of one or multiple conditions that must be met in order for numbers to be averaged.

Copying formulas

Formulas can be copied to adjacent cells using what Excel calls the “fill handle”:

If you place your mouse cursor over the fill handle, it will turn into a plus sign. If you now drag the fill handle, you can copy formulas to adjacent cells. Note that this will update relative cell references. For example, if you copy the formula `=A1+B1` from row 1 to row 2, the formula will become `=A2+B2`. Typically, this is the behaviour you want.

However, sometimes you might not want to update a cell reference. For example, you might have calculated the mean of some values and now you would like to calculate the deviation from the mean (i.e., subtract the mean from your other values.)

	A	B	C
1		response.rt	deviation_from_mean
2		634	11.75
3		549	-73.25
4		583	-39.25
5		723	100.75
6	Mean	622.25	

This is when absolute cell references (see Section) are required. Use the following formula in cell `C2` to subtract the mean from all values: `=B2-$B$6` (where `$B$6` is the absolute reference to cell B6)

When you copy this formula downwards, only the first cell reference will be updated. That is, the formulas will become:

- Row 3: `=B3-$B$6`
- Row 4: `=B4-$B$6`, etc.

If you want to copy a formula to a new cell **without changing anything initially**, the following procedure typically works best:

- Select the cell with the to-be-copied formula.
- Select the formula in the formula bar (the field next to fx).
- Press `Cmd + C` or `Ctrl + C`.
- Press `Esc`. **(Don't forget this step!)**
- Click on the cell to which the formula should be copied.
- Press `Cmd + V` or `Ctrl + V`.

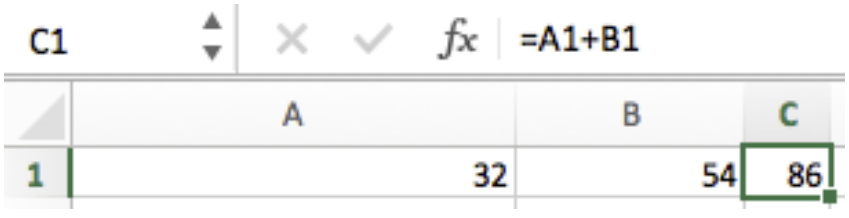
Self-study

Basic arithmetic operators

Below are some examples for formulas using arithmetic operators.

Add the values in cells `A1` and `B1`: `=A1+B1`

To get the answer in `C1`, you need to add the formula `=A1+B1` to the cell `C1`:



Subtract the value in `B1` from the value in `A1`: `=A1-B1`

Multiply `A1` with `B1`: `=A1*B1`

Divide `A1` by `B1`: `=A1/B1`

You get the idea. More information on arithmetic operators can be found here.

Other functions

Other potentially helpful functions are:

- `IFS`: test multiple conditions and return a value corresponding to the first TRUE result; unlike the IF function, IFS allows you to test more than one condition without nesting
- `COUNT`: count cells with numbers; will not count empty cells, logical values, text, or error values
- `COUNTA`: count cells with logical values, text, or error values; will not count empty cells
- `COUNTIF`: count cells dependent on the outcome of a logical test
- `COUNTIFS`: count cells dependent on the outcomes of multiple logical tests
- `ROWS`: returns the number of rows in a range
- `MEDIAN`: calculate the median
- `STDEV.S`: standard deviation of a sample
- `CORREL`: correlation

More details about all these function can be found in the Excel help and online.

Excel formulas and functions quiz

Self-study

This is a formative Excel quiz about content covered in Lab 9.

This section contains interactive content which is not available in the PDF version. Please visit the online version to see it.

January exam

Further information about the January exam can be found in the summative assessments overview in **?@sec-summative-assessments**.

Please note that the ExamSys practice exam is now available. There are 30 questions overall (about half the number of questions that are in the actual exam). If you would like to simulate the time constraints of the actual exam, you should give yourself 30 minutes to complete the practice exam.

You can repeat the practice exam as often as you would like to. If ExamSys tells you that "you will not be able to go back", it is referring to your current attempt, not future attempts.

Note that ExamSys allows you to grey out options that you are sure are incorrect. **To grey out an option right-click on the text of the option**. To bring it back to black, simply right-click again on it. Note: It is important to **always answer the question by left-clicking on one of the answers/radio buttons**. Greying out all options except one will not answer the question!

That said, it might be worth thinking about how often you can sensibly repeat the practice exam (where "sensibly" refers to whether or not the results are going to tell you anything about how your revision is going). You could just take a look at the questions initially (without answering them). This would give you an idea of the sort of questions that could come up (all the time keeping in mind that of course it won't be exactly these questions...). When you actually take the quiz, you don't want this to be too early (if you haven't revised the material yet, your quiz performance will not tell you much about your revision), but also not too late (clearly, you want to have some time left to continue revising after doing the quiz).

Other strategies are certainly possible. Which is perhaps a good opportunity to remind you of the article by Bjork et al. on learning effectively mentioned in **?@sec-note-taking**.

Explore, apply, reflect

Lab class/ Self-study

Main exercise

Goal: In this exercise, you will practice the Excel features introduced in the first part of the session. The aim is to reinforce your understanding by creating small example files or working with real data from PsychoPy output files.

Step 1: Choose your data

You have two options:

Create your own small dataset

Add reaction times (RTs) and accuracies from a small set of trials (5-10). Make sure some of these are correct (1) and some are incorrect (`0`).

Use a PsychoPy output file

- Open the file in Excel and save a copy (if you haven't already done so). Make sure there are at least two incorrect trials.
- Convert all RTs to milliseconds.

Step 2: Apply basic Excel skills

Work through the following tasks. If you already know some of these, focus on the ones that are new to you.

Formatting

- Automatically adjust column widths.
- Use conditional formatting (e.g., highlight reaction times above a certain threshold). (Ask an AI how to do this.)

Formulas and functions

- Calculate the **mean reaction time** using `=AVERAGE(range)`.
- Compute **accuracy percentage** using a formula like `=correct_trials / total_trials * 100` (use a function to find the correct numbers—don't just type them!).

Functions

- Copy all RTs from correct trials to a new column.

- Use `IF` and `AND` in a formula.

Sorting

- Sort trials by reaction time.
- Sort trials by accuracy, then by reaction time.

Step 3: Save your work

Save your Excel file on OneDrive with a clear name (e.g., `Excel_Practice_Lab9.xlsx`).

Challenge exercise

Your task is to explore one of the two datasets below in Excel and perform some basic analyses—look for patterns, trends, or anything interesting that stands out.

When you discover an insight: Add a sticky note to the respective Microsoft Whiteboard.

Briefly describe:

- Your question (what were you curious about?).
- What you did in Excel (e.g., the formulas or sorting/filtering approach you used).
- What you found out (your insight or observation).

Be creative and curious. There's no "right" answer! You are welcome to use Copilot (or other AIs) to help you find answers to any questions you might have.

The datasets and their whiteboards:

- **Amazon Echo Dot Reviews**: Contains customer reviews of the Echo Dot, including rating, review text, color, whether the purchase was verified, and review date. Amazon Reviews Whiteboard
- **Student Performance**: Includes information about students' study habits, attendance, favourite subject, learning method, smartphone usage, and exam scores. Student Performance Whiteboard

Click here to download the Echo Dot reviews.

Data Source: Amazon Echo Dot 2 Reviews Dataset by PromptCloudHQ is licensed under CC BY-SA 4.0.

Click here to download the student performance data.

Data Source: Student Academic Performance Dataset.

References