BASIC RULES FOR THE DEVELOPMENT OF SPREADSHEETS WITH MS EXCEL

1. Data Input-Output Model

The fundamental problem of a vast majority of Excel users is that they start badly from the beginning. I know it is a strong statement as a starter, but I suggest that we begin quickly with what is basic and important.

Our goal is amazingly simple: get the maximum performance with the least possible effort to be lazy…sorry, efficient! To achieve this goal, it is necessary to respect some rules that I will try to explain little by little. Most probably, you are already applying some of them; others will seem obvious or an unnecessary waste of time (by the way, these are the ones I fear the most).

In any case, I am going to try to prepare a small manual of good Excel practices that will help us get much more out of one of the most useful tools that exist - a tool that is sometimes very badly used. So, to avoid falling into *The Hell of Excel*, let's do it!

To become an Excel Warrior you will need to follow some commandments.

First Commandment:
I WILL NEVER EVER TYPE DATA INTO THE FORMULAS AREA

This is the first commandment because it is the most important of all, and one of the least respected. In my almost twenty years of teaching experience at IE Business School classrooms, I have found that this "mortal sin" is committed by all types of users: beginners, intermediate and those who think they are advanced users ... AND the most curious thing about this sin is that, very often, it shelters under the sad argument: "*Yes, I know, but I am in a hurry and I cannot waste time with unnecessary tricks ... If needed I will correct it later ...*" Tricks? Unnecessary? Later I will correct it? Not following this Commandment implies opening the doors of *The Hell of Excel*.

The starting key for the proper use of the spreadsheet is to use the Data Input-Output Model. This model is very easy:

The data must be separated from the formulas and the formulas from the data.

Our sheet must be defined with a specific area (or several areas) where we have the data (values, figures, percentages, product names, employee names, etc.) and another separated area (or areas) where we develop the formulas/model/processes that will provide us with the final results sought. In the cells that contain these formulas, **not a single piece of data should ever appear**. This data/value/number/figure/etc., must be allocated in the Input area, while **in the formulas we will always refer to the cell where the data is allocated**.

Here are some very simple examples that I think perfectly illustrate what should and should not be done:

Originally created by Professor Kiko Rial. Translated and adapted by Professor Guillermo de Haro. August 2020

	A	B
1		
2	**Price**	2.500,00 €
3	**VAT**	21%
4		
5	**GOOD**	**HELL OF EXCEL**
6	=B2*(1+B3)	=B2*1,21

Obviously, we can handle more complicated and/or extensive models than these simple cases – which is all the more reason to respect this commandment to the extreme! To see it in a practical way, consider the following situations in our second example:

		Year 1	Year 2
1	*Data Input*		
2	**Sales Year 1**	1.875,00 €	
3	**Sales Growth**	8%	
4	**COGS**	67%	
5			
6	*Data Output*		
7	**GOOD**	Year 1	Year 2
8	Revenues	=B2	=B8*(1+B3)
9	COGS	=B8*B4	=C8*B4
10	Gross Margin	=B8-B9	=C8-C9
11			
12	**HELL OF EXCEL**	Year 1	Year 2
13	Revenues	1875	=B2*(1+1,08)
14	COGS	=B8*67%	=C8*67%
15	Gross Margin	=B8-B9	=C8-C9

1. I have mistaken the sales figure for year 1 and need to change it.

If you have respected the first commandment you will only have to go to cell B2 and substitute the correct number in that cell. Job done in a couple of seconds. If you have entered the value 1,875 in one or more cells in the formulas area, what should you do? Indeed, look for each and every one of them, and modify them one by one (and don't forget any ...).

Again, this is a small Excel file, with a small example. If you have a file with a company budget, we can be talking about hundreds of megabytes. This is not efficient! You will become the Chief Searching-in-Excel Officer.

2. **I want to see how a reduction in the COGS to 55% would affect the gross margin of year 2.**

With the Data Input-Output Model correctly implemented, we just go to B10 and simply write the new value (55%). If, on the other hand, and to avoid wasting time and all that stuff, you have decided to write 67% within the formulas directly... bad news.

3. **I want to propose different scenarios for my sales estimate.**

If you have done it correctly, you can use the Scenarios Tool in a tremendously simple way, as we will see. Otherwise, you simply cannot use it. So, if you need scenarios you will probably choose to start inserting new sheets; copy and paste the model; substitute values; and so on (and this is not wasting time, of course ...).

We could continue providing countless everyday situations that show us, rather than the help friend that Excel is, *The Hell of Excel*.

This misuse of the spreadsheet is quite frequent. I have verified this both in the models that my students create, and in many of the consulting and executive jobs that I have done in several companies. The result of this misuse is hours and hours spent searching for errors; distrust in our own models (it is very typical to find Excel users with a calculator next to them to check that the results are correct); having to start a model all over again because we no longer know what it contemplates or what it calculates or how it does it; etc.

The solution to this situation (and to the unnecessary and inefficient hours lost) is very simple. Please, trust me, follow the first commandment and you will be rewarded with laziness...sorry, efficiency! And you will become an Excel Warrior.

2. Designing the Model

Continuing with our little manual of good practices in Excel, I will address the "second commandment" of our "Manual for Excel Warriors." The issue that we will deal with is not as tangible as many others that will concern us later, but I hope you know how to give it the importance it has, which, believe me, it is huge.

Second Commandment:

FIRST THINK, THEN OPEN EXCEL

Obvious, isn't it? Well, the truth is that this is the second terrible sin that many Excel users commit.

Before starting to develop any model in a spreadsheet, we must spend as much time as necessary to have a truly clear idea about where we are and where we want to go. Otherwise, and taking Murphy's law into account, we will end up making the least flexible approach that complicates our existence the most when we want to take advantage of our model.

Many experts recommend the so-called 80/20 technique, which means spending 80% of your time thinking and 20% "executing." Without wishing to create further discussions (and much less offend anyone) this theory seems to me just that: a theory. Any professional who uses spreadsheets daily knows that (among other things) it is very difficult, if not impossible, to estimate the time it will take to develop a model- Consequently, it is even more difficult to calculate 80% -20% of a figure that I do not know.

In any case, I would like us to truly internalize the underlying idea implicit in this theory: we must give real importance, both in quality and quantity, to the time provided to determine what I want to do and where in the Excel file I want to go. I know that many of you will think that sometimes you may not have time to think ... but if you read this last sentence again, you will realize that then something does not work.

As always, allow me to use a couple of elementary examples which clearly illustrate the problem we are talking about (and consider: if something this simple is already complicated, what is going to happen when we deal with something complex?)

a) Imagine that we want to estimate the evolution of our sales in the next years. Starting from the sales figure for year 1, we are going to apply a different percentage of sales growth for each year. We want the data output to be horizontal: that means having the "estimated sales" with the label "Revenues" in a row and the labels of the years, with their corresponding figures, in several columns:

		Year 1	Year 2	Year 3	Year 4	Year 5
10						
11		Year 1	Year 2	Year 3	Year 4	Year 5
12	Revenues					

If I start typing immediately, without even thinking for a second, the most normal thing is that I arrange the data entry in the following way:

	A	B
1		
2	Sales Year 1	600.000,00 €
3	Growth Year 1	5%
4	Growth Year 2	7%
5	Growth Year 3	11%
6	Growth Year 4	4%

With this approach, we will immediately notice that we will need to create 5 formulas to obtain our estimated revenues. Nevertheless, if we think a little bit more, we can see that with 2 formulas we can find these figures. I can check it by posing this other data entry:

1				
2	**Sales Year 1**	600.000,00 €		
3				
4		**Year 1**	**Year 2**	**Year 3**	**Year 4**
5	**Growths**	5%	7%	11%	4%

b) For the following example we are going to retrieve the data entry that we will raise in the Challenge 2, and that you can see below:

2	**Loan**	150.000 €
3	**Yearly IRR**	6%
4	**Years**	25
5	**Yearly Payments**	12
6	**Pre/Post**	0
7	**Final Value**	- €
8		
9		
10	**Fee**	

Note that in this Data Entry we contemplate, among others, the concept of "Yearly Payments" Why? Because although the most common payment frequency is monthly (12 payments per year), we might be interested in calculating the resulting installment for bimonthly, quarterly, semi-annual, or even annual payments. As we have committed to abiding by the second commandment, we have anticipated and previously thought about what analysis or information this model might require. And that is why performing this type of calculation will only involve the "tiresome" work of modifying the existing number in cell B5. Unfortunately (because they do not stop to think this through), most Excel users will have built the model by dividing and multiplying (depending on the argument in question) directly by 12.

Knowing everything that I will need from a model in advance is impossible. But there is quite a difference between knowing **everything** and spending a reasonable time asking ourselves crucial questions, such as:

- What information will I need

- How detailed will I need this information to be

- How often will I need it

- How many other users will be "manipulating" this model.

This analysis makes a significant difference. And believe me, experience shows that this will be time very well spent and that it will save you a lot of trouble and, of course, time.

3. Basic Excel Development Rules

Here I am going to provide several little tips that, if followed, will end up saving us a lot of time and trouble with the spreadsheet.

Third Commandment:

I WILL RESPECT THE BASIC RULES OF "CIRCULATION"

Once we have taken the time to think about the model that we are going to develop (see Commandment 2), we should be more or less clear about the structure and distribution of the Data Input and Output of that model (see Commandment 1). So, the time has come to get down to business.

From here onward, it is particularly important to be extremely structured and methodical, and it is also very important to follow what I call the **Basic Rules of "Circulation**." These are a series of small rules that apparently do not have much importance but, as you will see, they are **extremely important:**

1. **Differentiate the Inputs from the Formulas and the Data Outputs with colors**.

    We have already covered the importance of perfectly separating the Input from the Data Output (see the first commandment). Once this is done, I recommend that you use a color code to help you identify Input and Data easily and immediately. In this way, the user (or, in most cases, users) will know intuitively and clearly which cells can and cannot be manipulated. For example, use a blue and bold font for those cells that contain DATA (you can also highlight the labels of the data with a maroon background color and white text or orange background color and blue text, etc.). Keep the font color black and without filling for cells that contain formulas (the labels of these can be highlighted with a gray hatch, for example).

Loan	**150.000 €**
Yearly IRR	**6%**
Years	**25**
Yearly Payments	**12**
Pre/Post	**0**
Final Value	**- €**

Why? Some may talk about the KISS rule: **Keep It Simple, Stupid**. I prefer to use the "Grandma Rule:" if your Grandma understands it, so will your boss and your customer (not to mention your Excel teacher!)

When you create an Excel file, you know what you wanted it to do. But when the file is going to be use by other people - when it is going to "circulate" along the organization - the easier it is for others to understand what they have to do and what each element in the spreadsheet is, the better.

				Year 1	Year 2

Sales Year 1	1.250 €
Expenses Year 1	300 €
Sales Growth	10%
Expenses Growth	20%
COGS	72%

	Year 1	Year 2
Sales	1.250,00 €	1.375,00 €
COGS	900,00 €	990,00 €
Gross Margin	350,00 €	385,00 €
Expenses	300,00 €	360,00 €
Operating Profit	50,00 €	25,00 €
Accumulated OP	50,00 €	75,00 €

2. Leave the decoration for last.

A very common practice among spreadsheet users is to start a new model by dedicating a good deal of time to "decorating the sheet" with tasks such as changing cell alignments, filling it with a range of colors that they like, using a little bold around here and another bit of italics over there, combining a few cells, and, of course, using a good set of borders to highlight tables and important cells...

My advice: DON'T waste any more time on this at this stage. Dedicate yourself to formulating first. Then, when you have finished building your formulas, you can decorate your model. In this way, you will not waste time dragging –then correcting-unwanted formats (and deleting and redoing the formatting when you paste, rather than copy, the formatting) when you have to copy and paste formulas and insert rows and/or columns.

3. Do not change the initial alignment of the cells.

As is well known, and to reinforce the previous point, Excel allows us to quickly differentiate between a text input and a value just by observing where the input aligns:

- If the input aligns to the left when we enter a value, we will automatically know that it is a text.

- If it aligns to the right, then we will know that it is a number.

This small detail may seem trivial, but all help is useful in detecting possible errors in our models and, consequently, losing the minimum possible time correcting them.

To give an example of a very frequent error, try entering 4.45 and 4,45 in two different cells and you will see how Excel aligns this input differently:

	A
1	4,5
2	4.5

Furthermore, be careful with the notation you are using:

- In Europe the decimals are separated with commas, generally: that is why this example implies that 4,5 is considered a number.

- In the USA and many other countries, the decimals are separated with a dot, using the commas for the thousands.

- Countries like India use a completely different notation.

Consequently 2,500 will be two thousand and five hundred for an American, but two and a half for a European, who will be puzzled by the number of unnecessary zeroes after the comma.

4. Lowercase formula.

Another Excel "mini help", perhaps less known than the previous one, is that if we write the names of the functions and/or cells in lowercase correctly, when we press Enter after finishing our formula, it converts said name and references to uppercase. If it doesn´t, it will be because we have entered the wrong name of the function and/or the cells.

Try it out for yourself:

- Try to enter the formula =sum(f3:f8) in any cell and you will see that after pressing Enter, Excel transforms it into =SUM(F3:F8).

- Try to write, for example, =add(f3:f8) and you will see that will return the error #NAME? and that will keep the (add) function lowercase. If you do not use the "=" then the text will be seen directly in the cell as typed, add(f3:f8)

5. Use "Center Across Selection" instead of "Combine Cells."

Using "Combine Cells" is also a very widespread practice – unfortunately, as many of those you use it will tell you, it generates many later problems with formatting and selection of values.

A better option is to select the cells in which you want to center a label, click CTRL+1 (or COMMAND+! In a Mac), which will open the menu of the wizard to Format Cells. In the Wizard, go to Alignment/Horizontal and select "Center Across Selection." The final look will be exactly the same!

Try to use both approaches on a sheet to see what happens:

- You will see that, if you are using combined cells, when you click on them Excel selects B2 and C2 as a single cell

- When doing it with centered cells, you can click B4 and C4 independently.

This little trick will save you the insufferable message of "this operation requires that the cells to be combined have the same size"

5. Use Names instead of references

Whenever you can, use Names in formulas instead of the cells to which the range refers. There are several ways to name a cell or a range; the easiest is to select them and go to the editing box on the upper left area.
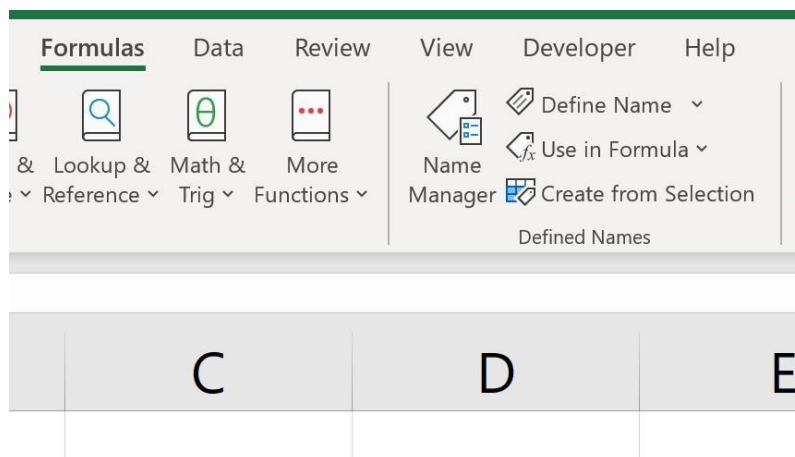
Let's see an example for A1. There we see that when we select the cell A1 we can see A1:
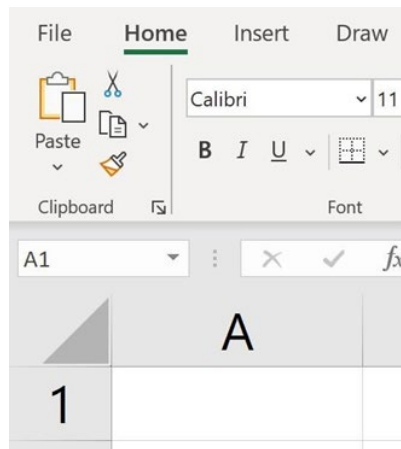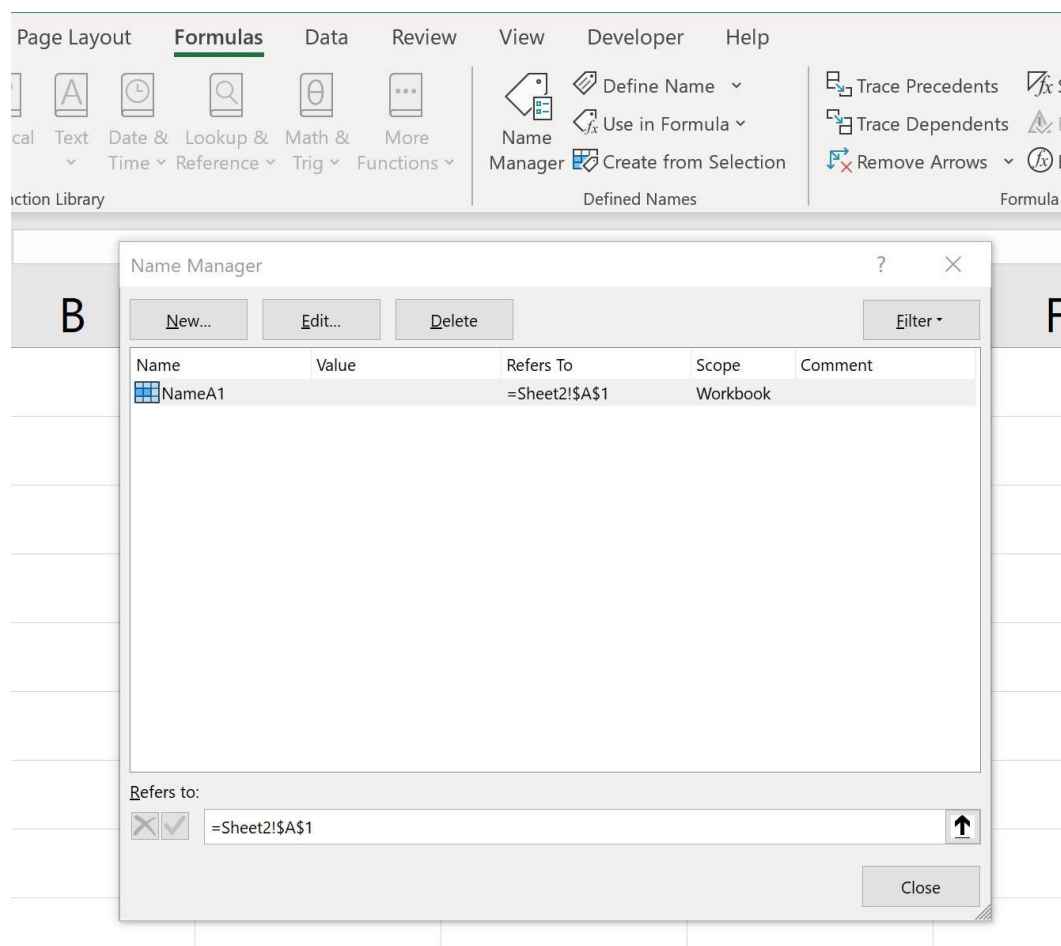
However, we can also name it. We cannot type most of the special characters there, nor can we add spaces. In this case we used NameA1 to name that cell:
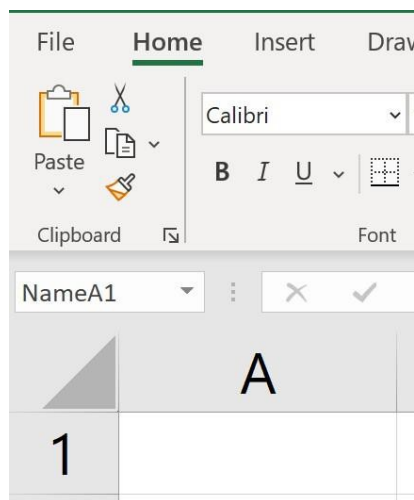
Now we can go the Name Manager in formulas to find the new name:

Now we can use the name to refer to the cell. This is called Natural Language. With this example it may not seem like a dramatically better way of doing things, but when we have a big file and we name variables like "Revenues_Year_1" or "Sales_Growth" it makes more sense, and it is easier to work with data this way.

This practice will allow you, among other things, to:

a) **Quickly locate the location of your tables.** If you give your tables a name you will only have to open the Name Box (it is the drop-down box that is to the left of the formula bar) and clicking on the name in question and it will automatically take you to the sheet (inside the book) and area where it is located.

b) **Make formulas more intuitive and intelligible.** Formula understanding, especially when they are large, is easier if we use range names instead of cell references. This also makes it easier to locate possible errors.

c) **Take advantage of names to optimize the formulation.** We will see this in practice in the videos.

6. Check each formula you develop as soon as you finish it.

This is a rule that I constantly insist on. Every time we create a formula, it is crucial that we "waste time" in checking that it works correctly. If we wait until the end to see if the 60 formulas that make up my model work correctly and one of them has a mistake, it is going to be a very considerable headache to detect where we have gone wrong (and if experience shows us something, it is that, no matter how much we know about this, they WILL be wrong).

7. Minimize the use of the keyboard.

The more you type, the more likely it is that you will make a mistake when formulating. A very silly example to understand is with the formula =A1+A2... I recommend that you avoid generate this formula by typing "A1" and "A2" with the keyboard. Instead, is it much better to click on the cells A1 and A2.

The same applies when working with ranges: select the range instead of typing the range in, for example, A1: A20.

8. Maintain uniform structures in your data outputs.

This will allow you to save many formulas since you can copy and paste them instead of having to develop them again.

A typical example is an income statement. In an income statement, the subtotals like Gross Margin or Operating Profit follow the same formulation: a figure minus another equals a subtotal. This will allow us to copy the formula in the first case (Gross Margin), and paste it directly into all the other similar ones (EBITDA, Pre-Tax Profit., Net Profit, etc.). With one formula, we end up solving for several cases.

If, on the other hand, you dedicate yourself to assembling seemingly more colorful models to leave spaces here and there so that, as I have heard on occasions, "you can see everything more clear and cleaner," then you will be forced to formulate many more times than necessary, since you will not be able to copy and paste as we explained in this example.