

Laboratorio Economia e Finanza



Marco Delogu

Data Management with Excel. Introduction to Derivatives

- ▶ What did we do yesterday?
- ▶ Outline of the course
- ▶ A brief recap of some of the most important Excel functions
- ▶ We discussed and saw a first example of a VBA code
- ▶ Now we deal with Data Management features with Excel
 - ▶ Notice, before doing any kind of analysis we need to put the data in the right form. Usually this a time demanding phase and sometimes neglected.. however no data in hand.. no analysis
 - ▶ Table, Filter and Pivot Table
 - ▶ Vlookup and Match Index
 - ▶ Record a Macro, our first VBA function
 - ▶ Write a VBA function
 - ▶ Introduction to Derivatives

- We start with a dataset that I have used in my research *“The role of fees in foreign education: evidence from Italy”*

1. A brief digression

- 1.1 This paper employs data on foreign students inflows to Italian universities.
 - 1.2 This research paper finds evidence that increasing the level of fees actually reduces the flows of foreign students... Looks like an obvious result.. but previous research relying on data retrieved at the country level could not find such evidence

2. What we do today... we recap or see for the first time some very useful features of Microsoft Excel

- 2.1 Notice that the data is not exactly the one that I used, however it keeps the same structure so.. you cannot replicate the regression of the paper.. not the subject of the lecture!
 - 2.2 We use this dataset, **“datamanagement.xlsx”**, look at the data saved in the sheet labeled “data”

3. **datalecture2.xlsx**

- 3.1 Contains 13495 rows and 10 columns
 - 3.2 Each rows report the number of students coming from one country and going to a specific university. The other columns report some information specific to the university
 - 3.3 The file contains a summary sheet that explains what is contained in each sheet (look at the file in the moodle)

- ▶ **Institution:** name of the university
- ▶ **Countryoforigin:** name of the country of origin
- ▶ **Foreignstudents:** number of students, this one is not the actual data)
- ▶ **Hostcapacity;** total number of students enrolled at the university, at a given year
- ▶ **StataleLibera:** type of the university, *L* for private, *S* for public
- ▶ **Costofliving:** index of the cost of living, taken from the Numbeo website
- ▶ **Citygdp:** city for which the data on GDP has been taken
- ▶ **GDPercap2011:** value of average gdp per capita at this region, taken from Eurostat
- ▶ **Fees24h:** a proxy of the average level of fees charged by the University, data from Sole24ore, taken from Miur
- ▶ **Dummyenglish:** takes value of 1 whether the university was teaching an english course this year or not

Our data contains a lot of data.. sometimes we want just focus on a specific information. Excel comes to help us with some specific tools:

- ▶ **Table:** puts a filter in each column, and to select the type of data that we need
- ▶ **Filter:** similar to Table, however put the filter only where we want
- ▶ **VlookUP:** allows to merge/connect information defining a column as the identifier
- ▶ **MatchIndex:** allows to merge information using two identifiers
- ▶ **PivotTable:** allows to create information and retrieve easily information. It produces graphs quickly

Table allows us to to easily grasp some key features of the data.

- ▶ Select your data and go to Insert→Table or press CTRL-T
- ▶ What happens once the data is in Table format? You can select only some specific data that you want to.
- ▶ What if you want to get rid of the Table?
- ▶ Select some cells. Go to Table and select “*Convert to Range*”...notice that by doing so you get your data back. However the formatting of the Table is kept. You can go back easily to the basic formatting

- ▶ Imagine that we want to see only the data of a specific university...Once you have the data in the Table format it suffices to select it
- ▶ Or imagine that we want to see only the data for Private university.. from the column select only the private universities
- ▶ Examples:
 - ▶ Select the data of Uniss (sheet example1)
 - ▶ Select the data where StataleLibera value equal S (sheet example2)

! This is very important, wherever you will go to work you may be probably faced with extremely big excel sheet. It is quite likely that you will be asked to retrieve a specific information. By transforming the data in the Table you can easily make yourself and, more importantly, your boss happy

- ▶ Our data contains information concerning the same universities several times... Let's say that we were asked, to produce me excel sheet containing only a list of the universities.
- ▶ You may try to get the job using Table... but this can be risky. Better, Excel has available a command that allows you to easily obtain this information
- ▶ Select the column with the header institution... go to Data and press, close to Filter the icon called Advanced...Here select Copy to Another Location... in Copy to Select Cell L1
- ▶ Drawback of Filter: can copy only in the active sheet...
- ▶ In addition to the names we now know how many universities are present in our database. It contains information for 78 universities.
- ▶ We can do the same for the origin countries listed in the database

- ▶ Select column B
- ▶ Go to the Data, and close to Filter click Advanced.
- ▶ Select “Copy Unique Records Only”
- ▶ Select Copy to and indicate cell M1
- ▶ Here you find the list of countries... Now we have the list of countries and discover that the number of origin countries is equal to 173!

- ▶ In my mind is the most useful function of EXCEL... **VLOOKUP**...V stands for Vertical

Syntax: Vlookup (value, table, col_index, [range_lookup])

- ▶ Arguments:
 - ▶ **value**: the value to look in the first column of the table
 - ▶ **table**: the table from which to retrieve a value
 - ▶ **col_index**: the column in the table from which to retrieve a value
 - ▶ **range_lookup**= TRUE=Approximate match, False =exact Match (If the value is not found in the data range, you will get an error value #N/A)
 - ▶ remind that in this case it can be useful to use the function **iferror**...think about it!
- ▶ First I will focus with examples requiring the Exact Match. To use the Approximate match the data must be sorted previously.

- ▶ Imagine that someone knows that you have this data on foreign students and ask you a specific information
- ▶ They ask you to know whether the university *Sassari*, *Siena* and *Insubria* where providing a course in English. Remind that in the dataset this information is stored in column *J*, labeled *dummyenglish...*
- ▶ You may try to solve the problem in two ways: (1) can retrieve this information manually just scrolling the data.. (2) or using table...
- ▶ However... much better to use Vlookup... why? you can almost automatize your answer.. and you will be able to deal with possible additional inquiries.. (and in the real world this happens quite often, namely one second you got the answer.. they will ask you a close and similar question!)

- ▶ Create a new excel sheet, name it *Vlookup1*,
- ▶ Label the first column, cell A1, with “Institution”
- ▶ write below, Sassari (A2), Siena (A3) and Insubria(A4)
- ▶ In cell b1 write English-course. We want to report, here, whether the university was providing a course taught in English or not.

How to do?

- ▶ You can select the data using the mouse...but it can make you nervous, much better use the combination CTRL SHIFT → ↓
- ▶ Use the function Vlookup....then you must have

$$= vlookup(a2, data!A1 : j13495, 10, FALSE)$$

- ▶ Notice the dollars.. why are they necessary? for the problem in hand we must keep the range fixed (always look in all dataset). (*absolute references!*)
- ▶ the first argument is the identifier , the value to look for in the first column of the table
- ▶ the third argument identifies the column from which we want to retrieve the data

- ▶ Sometimes (but it never happened to me!) you may be interested to look for an approximate value...
- ▶ Notice that the data must be sorted previously
- ▶ If you do not sort the data previously... the formula does not work
- ▶ To sort the data go to Data and select Sort
- ▶ Look at the worksheet Vlookup2

- ▶ In my mind is the most useful excel function. Allows to retrieve information promptly and easily
- ▶ Important: In the real world you receive data from different sources. but somehow these data are linked and can be combined. Vlookup allows you to exploit the links to eventually merge the datasets
- ▶ **Main Drawback:** unfortunately you can match the data only by **one identifier...**
- ▶ But sometimes we need to merge data according to multiple criteria...
- ▶ In this case, Vlookup cannot help to solve the problem... however you can overcome the issue using jointly two functions: *Index* and *Match*
- ▶ When first study them separately

Match: a function that returns the location of a value in an array

Syntax: *Match (lookup_value, lookup_data, [match_type])*

- ▶ *lookup_value*: the value that you are looking for
- ▶ *lookup_data*: is the array of values which contains the data that you are trying to find
- ▶ *match_type*: defines the type of match that you are looking for, put it to 0 to look for the exact match. If you do not put anything (or put 1) the values in the lookup array must be sorted in an ascending order. Then Match will give you the value that is lower or equal to the lookup_value. The other value that is argument can take is equal to -1. In this case the values in the lookup array must be sorted in descending order. Then Match gives you the value that is greater or equal than the lookup_value
- ▶ look at the excel sheet Match Index... as you can see you obtain a number, which is the row number (relative number) where, for the first time, is found a value equal to the value that we are looking for (in our example is Sassari)... important: do not forget to put absolute reference for the cells

Index function: takes as input the location and return you the value that the cell contains

Syntax: ***=Index(array, row_num, [col_num])***

- ▶ ***array:*** is the table of data that contains the cell that you want
- ▶ ***row_num:*** is the relative row number of the cell you want
- ▶ ***col_num:*** is the relative column number of the cell that you want
- ▶ Remind that match gives you the relative cell number... then combining index and match you can obtain... the same result that you would have obtained by using Vlookup

- ▶ We notice that using jointly Match Index we obtain the same result that we obtained using only Vlookup
- ▶ *Let's do it* Consider the formula below:

INDEX(data!\$A\$2:\$F\$13495,**MATCH**('Match Index'!A3,data!\$A\$2:\$A\$13495,0),5)

- ▶ You must just fill the row_num argument of the INDEX function with the Match Function.
- ▶ What is Excel doing?: (1) the Match function retrieves the location that we are looking for. (2) The output of the Match index becomes an argument of the Index function. Likewise the Vlookup function Excel returns us the data that we were looking for.

- ▶ Another task..Now we want to look for multiple criteria!
- ▶ To do so we must convert our formula to be an *array formula* and learn a new excel feature.
- ▶ Consider the sheet labeled Match Index 2..Now we want to retrieve the number of Russian Students studying in Verona!
- ▶ To solve this problem we need to look for two criteria:
 1. *first criteria*: the university
 2. *second criteria*: the country of origin

Before our match index was already feed with an array..Now
Excel builds is own array from scratch!

- ▶ Before we used a Match formula that looked in this way:
- ▶ `MATCH(A2,data!A2:A13495,0)....` now it is a bit different, Excel, is taking advantage, but not showing to you, of logical operators
- ▶ To properly understand what is going on it is better to look directly at the formula:

$MATCH(1,(data!A2:A13495='Match\ Index\ 2'!A2)*(data!B2:B13495='Match\ Index\ 2'!B2),0)$

- ▶ What is the above function doing?.... it is looking for a value equal to 1....
- ▶ The formula now builds an array from scratch, first it look in the column A for a value equal to the first criteria.. then it looks in the column B for a value equal to the second criteria
- ▶ Once both Boolean values are equal to 1 it stops!... and the function returns us the relative position that we were looking for.....
- ▶ It is easy to see that now it suffices to put this value into the Index function to report the number of students, the value that we wanted
- ▶ Notice that you can put as many criteria you want!!!
- ▶ Important: this is an example of an array formula. Once you wrote it to obtain the result you must press CTRL SHIFT ENTER (and see the curly braces, `{formula}`) The combination is the same whether you work with Mac or Windows

- ▶ Summary of the example (see excel sheet Match Index 2)
- ▶ Define the criteria (in our case they are):
 - ▶ the name of the university
 - ▶ the name of origin country
- ▶ Use the match function to determine the position in the array (the one that contains our data) of the information that we are looking for
- ▶ Once we know the position we can use the index function to retrieve the information that we wanted

$$=INDEX(data!A2:F13495, \\$$
$$MATCH(1,(data!A2:A13495='Match Index$$
$$2'!A2)*(data!B2:B13495='Match Index 2'!B2),0),3)$$

- ▶ Notice that you need to press **CTRL+SHIFT+ENTER** after the cell is filled with the formula

- ▶ A Pivot Table allows us to retrieve information, very quickly...
- ▶ Go to the sheet Data
- ▶ Go to Insert and look for Pivot Chart, here click Pivot Chart & Pivot Table
- ▶ Finally choose new worksheet
- ▶ You have generated a new worksheet. You can label this sheet as you wish (give it a name) I opt to call it Data

- ▶ A lot of things can be done here....
- ▶ Imagine that we want to know how many foreign students were enrolled at the university of Sassari
- ▶ Let the Institution to become your filter
- ▶ Put in the section values *foreignstudents*
- ▶ Then we obtain a graph and the result... you can easily switch!
- ▶ As you can realize this command makes easily to recover quickly some kind of information

Record a Macro in Excel. See Excel File "IntroductionMacroVBA.xlsm"

- ▶ Very useful tool to automatize some boring task..

Take the excel file "" and consider the sheets: *Example 1*, *Example 2* and *Example 3*. In all of them

- ▶ A1="Nome", A2="Cognome", A3=ETA'
 - ▶ This is just a quick example... imagine that you have plenty of sheets that are similar and for some weird reasons you are asked to perform the same operation in all of them (think about formatting purposes)
- ▶ Also assume that you don't know how to program VBA, and that you are quite lazy.. However, your boss called you and he wants the A1 A2 appear in red and A3 in blue...
- ▶ You can do it manually... however, by doing in this way you'll spend all the afternoon

You can Record a Macro in Excel in two ways

- ▶ **Absolute References:** it will run the macro considering the same cells that you considered when you recorded it... the result of the macro (once you run it) does not depend on the active cell
- ▶ **Relative References:** it will run the macro in relative terms.. Easier to explain through an example. Let assume that you recorded something starting from A1 and then moving to D1. Then if you run the macro, and your active cell at the beginning is A2 it will do something to cell D2.. Once you using *Relative References the results depends on the cell that is active at the moment you run the macro*
- ▶ Notice that once EXCEL is recording the macro it allows, afterwards, to retrieve the code.

- ▶ Go to excel sheet “Example 1”. Then make cell A1 your active Cell
- ▶ Go to Developer, and then on the left of Macro you have the button Record Macro
- ▶ Give a name to the macro, in our case we call it *Example 1*
- ▶ Press it
- ▶ Then make cell a1 and b1 red and cell c1 blue
- ▶ then go to stop recording
- ▶ Now you have a Macro!
- ▶ Go to the sheet Example2 . Make cell A1 the active cell. Then go to Macro. Now a window should open, click Example 1 and then Run!
- ▶ Magic!.. Do the same for Excel sheet Example 3
- ▶ However this macro does not work well for the sheet Example 4.... However we can overcome the issue recording the macro using relative and not absolute references!

- ▶ Go to excel sheet “Example 1”. Then make cell A1 your active Cell
- ▶ Go to Developer, and then on the left of Macro you have the button Record Macro
- ▶ Give a name to the macro, in our case Example 1
- ▶ Press it
- ▶ Then make cell a1 and b1 red and cell c1 blue
- ▶ then go to stop recording
- ▶ Now you have a Macro!
- ▶ Go to the sheet Example2 . Make cell A1 the active cell. Then go to Macro. Now a window should open, click Example 1 and then Run!
- ▶ Magic!.. Do the same for Excel sheet Example 3
- ▶ However this macro does not work well for the sheet Example 4.... However we can overcome the issue recording the macro in relative terms

- ▶ Go to excel sheet Example 1, let's make it again all cells black (manually)
- ▶ Do the same steps of before, (namely go to record macro, but below click to the button *"Use Relative References"*)
- ▶ Save the Macro with the name Example2
- ▶ Depending on the problem in hand you can decide to record to record the macro in:
 - ▶ absolute references
 - ▶ relative references
- ▶ Now go to Developer, and Macro. Now You should see in the window both Example1 and Example2
- ▶ Go to Edit... you go to the VBA editor... notice that here you can change the code!
- ▶ Let's do it... replace
- ▶ Sub Example1() ' ' Example1 Macro '
- ▶ ' Range(A1:B1). with Range(b2:c2). and Range(c1). with Range(D1)....this one is a trivial example, but makes you understand that by understanding the code we can do a lot of things thanks to VBA
- ▶ Interestingly: recording a macro can be a good way to save time but also to learn some VBA syntax!

- ▶ Function:
 - ▶ The difference between a function and a sub in Excel VBA is that a function can (must) return a value while a sub cannot (just run a set of instructions). Functions and subs become very useful as program size increases.
- ▶ Notice that a function is something that you already know well, once you use EXCEL you take advantage of many of them
- ▶ Today we learn on how to build our own function in EXCEL
- ▶ Let's build our own function that computes for us the area of the rectangle
- ▶ Then this function will appear among the other existing EXCEL functions (such as LEN, Vlookup) and you can use it as many time as you want
- ▶ Just a quick reminder.... $AREA = b * a$ where b (length, in Italian base) and a (width, altezza in Italian) must be to positive numbers.

Let's build our first function.. we want to calculate the rectangle area. We need to open the Developer.
then click Visual Basic, or press ALT F11
Here we need to write our first function...
Our first function could not be easier than this one:

Function **area** (length As Double, width As Double)

area = length * width

End Function

Notice that we need to define the arguments (length, and width)
define the type... Function must return a value...area in our case
Check, now this function appears along the built in Excel
function!

Other thing.... To check that both the width and the length are both positive values we take advantage of the *If then else*
Notice that in fact you can compute the area of the triangle if and only if both the width and the length are positive!
Handling this situation requires to introduce the if then else statement

Function area1(length As Double, width As Double)

If ((length <= 0) **Or** (width <= 0))

Then area1 = errore

Else area1 = length * width

End If

End Function



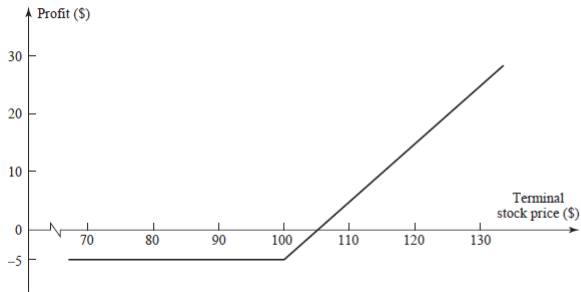
(Adapted from Wilmott's Books) *"The simplest option gives the holder the right to trade in the future at a previously agreed price but takes away the obligation."* Definition: **A call option is the right to buy a particular asset for an agreed amount at a specified time in the future**

- ▶ **strike price or exercise:** price at which we can buy the stock
 - ▶ **expiry or expiration date:** date at which we may exercise the option
 - ▶ **underlying asset:** the stock on which the option is based
- Denoting by S the value of the value of the underlying

$$\max(S - E, 0)$$

Figure taken from Hull, ed.8

Figure 9.1 Profit from buying a European call option on one share of a stock. Option price = \$5; strike price = \$100.



A put option is the right to sell a particular asset for an agreed amount at a specified time in the future

$$\max(E - S, 0)$$

Calls and puts are the two simplest forms of option. For this reason they are often referred to as *vanilla*



Figure taken from Hull, ed.8

Figure 9.2 Profit from buying a European put option on one share of a stock. Option price = \$7; strike price = \$70.

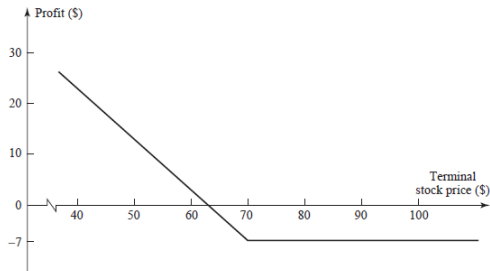
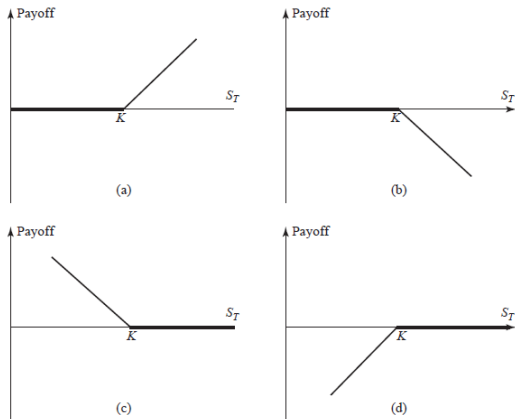


Figure taken from Hull, ed.8

Figure 9.5 Payoffs from positions in European options: (a) long call; (b) short call; (c) long put; (d) short put. Strike price = K ; price of asset at maturity = S_T .



- ▶ **Premium:** The amount paid for the contract initially
- ▶ **Intrinsic value:** The payoff that would be received if the underlying is at its current level when the option expires.
- ▶ **Time value:** Any value that the option has above its intrinsic value. The uncertainty surrounding the future value of the underlying asset means that the option value is generally different from the intrinsic value.
- ▶ **In the money:** An option with positive intrinsic value. A call option when the asset price is above the strike, a put option when the asset price is below the strike.
- ▶ **Out of the money:** An option with no intrinsic value, only time value. A call option when the asset price is below the strike, a put option when the asset price is above the strike.



- ▶ **At the money:** A call or put with a strike that is close to the current asset level.
- ▶ **Long position:** A positive amount of a quantity, or a positive exposure to a quantity.
- ▶ **Short position:** A negative amount of a quantity, or a negative exposure to a quantity. Many assets can be sold short, with some constraints on the length of time before they must be bought back.
- ▶ The writer of an option is the person who promises to deliver the underlying asset, if the option is a call, or buy it, if the option is a put. The writer is the person who receives the premium.

- ▶ Our previous definitions assumes that we can exercise the option only at the expiry date. This kind of options are known as **European Options**
- ▶ **American Options** allow the holder to exercise at any time before expiry. American options give the holder more rights than their European equivalent and can therefore be more valuable, and they can never be less valuable. When you hold an American Option what is important is when to exercise
- ▶ **Bermudan options.** These kind of options allow to exercise on specified dates, or in specified periods. In a sense they are half way between European and American since exercise is allowed on some days and not on others.

We can now write some VBA function that will compute for US the payoff function at the expiration date of a European-Call/Put Remind

Call

$$\max(S - E, 0)$$

Put

$$\max(E - S, 0)$$

Function PAYOFFC(final_value, strike)

 If (final_value > strike) Then

 PAYOFFC = final_value - strike

 Else PAYOFFC = 0

 End If

End Function

Function PAYOFFP(final_value, strike)

 If (final_value < strike) Then

 PAYOFFP = strike - final_value

 Else PAYOFFP = 0

 End If

End Function

- ▶ Writing our function we did not include the premium that the individual has to buy up front for either a Call or a Put. **Homework:** Adapt the VBA function to take into account the option price when computing the Payoff.
- ▶ *Technicality:* Also, we should consider that the premium is paid up front and the profits, if any, are received only at the expiry date. Consequently we should either discount the profits or multiply the premium by $e(r(T - t))$, to evaluate everything at the expiry date.

Today is 1st April and the Asset *LaboratorioLecture* is worth 650. In the market you can buy a call option with strike price 680 for the 1st May at the price of 30. You expect that the stock rises during this month. How you can profit? Namely consider that the price of the stock the 1st of august is equal to 720. Determine the profits that you'll get, both in absolute and percentage terms, for the two strategies below

- ▶ buy the asset at 650 the 1st of April
- ▶ buy the call option at 30 and exercise it. (take advantage of your own *VBA* function)



- ▶ This is an example of gearing or leverage. The out-of-the-money option has a high gearing, a possible high payoff for a small investment. The downside of this leverage is that the call option is more likely than not to expire completely worthless and you will lose all of your investment (What is the return of your investment in such a case?)
- ▶ Highly leveraged contracts are very risky for the writer of the option. The buyer is only risking a small amount; although he is very likely to lose, the downside is limited by the initial premium!