M6 Excel Topics

Topics in Insurance, Risk, and Finance [1]

Professor Benjamin Avanzi

21 August 2023

[1] References: Slager and Slager (2020) and Katz (2023) | → latest slides

THE UNIVERSITY OF MELBOURNE

Why Excel?

- Excel is great to spread data and calculations in a tabular form, and have a visual overview.
- Most financial modelling is done in Excel (at least initially).
- In actuarial work, many more advanced codes (in R, Python, C++, C#, VBA, ...) often starts with someone playing around in Excel, and once proof of concept is approved, this moves to proper coding.
- It is assessed in CM2-B (!).

1 General Considerations

- Why Excel?
- Issues with Excel

Issues with Excel

Excel is notoriously problematic in certain areas:

- Lack of transparency - one can't see the formulas unless you click in a cell; alternatively, you can't see the formulas and understand where the numbers come from unless you are in Excel (a problem for reports, presentations etc).
- Mistakes can be tiny but have huge consequences in the end (the job of Excel auditor actually exists!).
- Lack of good documentation capability (as opposed to code); this makes collaboration and audit difficult, and creates an operational risk (e.g. builder leaves).

- Lack of rigour in the construction of a model (input, assumptions, intermediary calculations, output).
- Can't handle (seriously) large data sets.
- Lack of good and flexible data cleaning and manipulation capabilities.
- Sometimes code is a lot easier (e.g. flip a vector around, sum over a diagonal, . . .).

I know there are counter arguments for all of those, but this presupposes you know what the solutions are. You'll learn some of those here!

Assumed knowledge

See `prerequisite knowledge on the website`. Some extracts:

- Autofill: Chapter 3, p. 105-116, and p. 298-301
- Named ranges and constants: Chapter 7, page 312-332
- Absolute and mixed cell references ($): Chapter 7, pages 332-342
- New Excel 2019 functions (IFS, MAXIFS, MINIFS): Chapter 8, pages 381-398
- Formula Auditing: Chapter 9, p. 436-439
- Paste special (incl, e.g. `Transpose`): Chapter 11, pages 518-530

Page references are for Slager and Slager (2020), see `link here`.

Also, see tab `Prerequisite` in the `module 6 spreadsheet`.

3 General etiquette and tools

- Etiquette
- Principles

Etiquette

You should build your spreadsheet with (at least) the following **objectives** in mind:

1. so as to minimise chances or error (accuracy);
2. so as to minimise unnecessary calculations (efficiency);
3. so as to make the structure as clear as possible (transparency);
4. so as to make updates, changes and extensions possible and easy (extendibility);
5. so as to allow someone else to use it easily (user friendliness);
6. so as to allow someone else to verify it easily (auditability).

Those are, of course, interconnected. We could add more (for instance, automation of data input via an API, automation of communication objects such as charts, etc. . .).

3. General etiquette and tools
   - Etiquette
   - Principles

Principles

There is no single way to achieve the objectives above, but there are a number of **principles** that one could list, and that will contribute to meeting those objectives:

- Have a separate tab that collects all your assumptions that are valid for the whole spreadsheet (1, 3, 4, 5, 6).
  - Include some explanations about the source/justification of those assumptions.
  - Give your assumptions names (for instance, the technical rate of interest to calculate life insurance could be called `techint` or similar, for ease of later reference, and to make formulas more easily readable).
- Also include your data sets in separate tabs (1, 3, 4, 5, 6).
  - Name your data (including columns and/or rows if possible; e.g. `FIFA WWC` in the spreadsheet; see also "Named ranges and constants": Chapter 7, page 312-332).

- Consider colouring / contouring input and output differently (3, 5)
    - This is not always advisable, but if you have a large model with relatively few input (e.g. purchase price and interest rate for a property mortgage schedule) and/or few outputs (e.g. NPV) then this achieves many of the objectives.
- Use named ranges and variables as much as possible (1, 3, 5, 6), unless the variable is going to be used only once.
- Use more advanced formulas if shorter, and avoid too much nesting (1, 3, 4, 5, 6).

4 Tools for auditing

- Dependents and precendents
- Show formulas

Dependents and precendents

Use of dependents / precedents, e.g.:

1. go to tab `Dependents - Precedents - PPCI`;
2. click on one of the outstanding loss projected amounts;
3. make sure the formula tool tab is live;
4. click on the "trace precedents" sequentially.

This will highlight dependence of each cell to previous cells, sequentially, with arrows.

This is useful for

- understanding the structure of a spreadsheet;
- check formulas (audit);
- debug issues.

(Formula Auditing: Chapter 9, p. 436-439)

4 Tools for auditing

- Dependents and precendents
- Show formulas

Show formulas

- The “Formulas” tool tab should have a “Show formula” tile. This will replace all numeric values by formulas.
  - This is helpful to check what numbers are hard coded, and which are results of calculations. Together with dependents, it helps seeing if everything is as dynamic as it should.
- If you want a formula to be shown all the time start with an apostrophe `'`, and the formula will show as text.
- Note also the `FORMULATEXT()` formula which is a dynamic array formula requiring spilling (see later section “Dynamic Arrays”!).

(Formula Auditing: Chapter 9, p. 436-439)

5 Tools for analysing data

- Pivot tables
- Pivot charts

Pivot tables

- Pivot tables are often considered as very difficult to master, but they are not that difficult to start with.
- Example (in `module 6 spreadsheet`): FIFA WWC
    - Insert / Pivot Table.
    - See how you can use variables as filters, rows, or columns. Move them around.
    - See how columns can display other things than `Sum`, such as `Count`, `Average`, `Max`, `Min`, or `Product`
- Note there are recommended Pivot Tables (automated recommendation within Excel); in the case of FIFA WWC it is not very helpful.
- Reference: Chapter 15 of Slager and Slager (2020).

5 Tools for analysing data

- Pivot tables
- Pivot charts

Pivot charts

- A pivot chart can be created from the Pivot Table but also directly from the data.
- A major difference with start charts is that you it will be somewhat "interactive" - there will be buttons you can use to alter the chart.
- Example (in `module 6 spreadsheet`): FIFA WWC
  - Insert / Pivot Chart;
  - In the example I changed the style of graph to "Combo" to allow for the two different scales;
  - I also included a "slicer" (click on chart / insert slicer), in order to easily filter by squad.
- Reference: Chapter 15 of Slager and Slager (2020).

6 Dynamic Arrays

- Spilling
- Think in vectors
- The # sign
- New formulas

Spilling

- One advantage of programs like R are the easy use and manipulation of vectors.
- Excel can do similar things, and the vectors are called arrays. This is new (post Office 365), and is a bit of a game changer.
- Before Office 365, Excel was incapable of depositing results beyond just 1 cell. This is called “spilling”.
- Note Excell will need the required space to spill.
- Main reference is Katz (2023) - we’ll only introduce this here.

Example

- Here we introduce array formulas.
  - If you calculate the sum of an array you'll get a single number.
  - The result of an array formula (such as `LEN()`), when you input an array, will give you an array.
- See `module 6 spreadsheet`:
  - `LEN()` gives an array.
  - You could then get the sum without having to put the array anywhere: `SUM(LEN(B3:B6))`.
  - Note that when I wrote the above, it became `SUM(LEN(B3#))` automatically - more on that later.
- `SUM(LEN(B3:B6))` will work in any version of Excel because it requires only one cell to output, but not `LEN(B3:B6)` as it requires several cells ("spilling").
- Note you can spill named ranges, too!

6 Dynamic Arrays

- Spilling
- Think in vectors
- The # sign
- New formulas

Think in vectors

- Once you understand you can create vectors and either display or manipulate them, Excel becomes a lot more powerful.
- You can also use arrays in arguments of known formulas such as `VLOOKUP()`,
  - For instance `VLOOKUP(.,.,.,{2,5})` will return value from the 2nd and 5th columns row-wise.
  - If you use `VLOOKUP(.,.,.,{2;5})` (with the semicolon) they will display columnwise.
  - See examples in `module 6 spreadsheet`.

6 Dynamic Arrays

- Spilling
- Think in vectors
- The # sign
- New formulas

The # sign

- The # sign when added to a reference to a cell where a dynamic array is written will duplicate that array (and spilled results).
- It is shorter, but also it will dynamically change the size of the array
  - This can be desired or not.
  - See example in `module 6 spreadsheet`.

6 Dynamic Arrays

- Spilling
- Think in vectors
- The # sign
- New formulas

New formulas

- There are a number of new formulas which were available in coding languages like `R` for a long time, which can be useful, and which are now available in Excel
  - for instance `SEQUENCE()`, `UNIQUE()`, `FILTER()`, `RANDARRAY()`...
- Some of those are exemplifed in `module 6 spreadsheet`.
- You are encouraged to browse through. They really bring data handling in Excel a little closer to coded languages such as `R`.

Next steps

- All of Katz (2023) is relevant, but take it as a cook book for the assignment. You can go as far as you wish.
- Chapters 16, 17, 18, and 19 of Slager and Slager (2020) are out of scope.
- However, macros and VBA (which is Chapter 19) are essential components of Excel
    - I strongly encourage you to get started. Start by recording a macro, then play around with the code.
    - VBA allows more efficient calculations via compiled code, and is a powerful addition to Excel.
- Fun fact: the 2021 `Excel World Champion` is an actuary: Andrew Ngai, now Director at Taylor Fry.

References I

Katz, A. I. 2023. *Up up and Array! Dynamic Array Formulas for Excel 365 and Beyond*. Apress.

Slager, D., and A. Slager. 2020. *Essential Excel 2019*. 2nd ed. Apress.