

USE OF EXCEL FOR DATA HANDLING IN LABORATORIES

This Cook Book is not intended to give comprehensive guidance on every thinkable use of Excel in laboratories but it will point out some important pitfalls and some good advices to avoid them – both of which will require further and more detailed studies to give the full understanding.

Areas of application - *lots of exciting and challenging possibilities, but be careful*

Calculations

Excel is used in most laboratories as the primary calculation tool – especially for those calculations carried out routinely and built into fixed application templates (unfortunately often with the user not having the faintest idea of what is really going on). It is important to use the many functions and facilities in Excel (incl. more advanced calculations by help of macros) *in a conscious, systematic and controlled way!*

Statistics

One special group of calculations which can easily be performed by Excel are the statistical estimations and tests etc. The risk is though that people just pick a tool from the functions toolbox without knowing whether it is the right one for the purpose and without the right understanding for evaluating the results correctly (unfortunately the HELP facilities for these functions are not very well elaborated or really helpful!)

Use of Excel for statistics therefore requires a basic understanding of the statistical theory and prerequisites!

Data analysis

Excel is created to handle and (re)arrange big amounts of data, not only for basic and statistical calculations but also for organizing and categorizing the data to highlight any coherence between data, trends or effect of influence factors. The ultimate tool for that purpose in Excel are the Pivot Tables and related charts but as always *a good analysis takes some good and careful planning of the whole set up before entering the data.*

Illustration

Excel has at least about 100 different chart types to help you illustrate your data and accentuate any possible coherence between sets of data – and which can easily be set up by using the built-in Wizard. Most of the options are though normally not relevant for illustrating laboratory data, so *be sure to choose and set up the right chart (e.g. a calibration curve) for the purpose which truly represents the actual data and doesn't lead to misinterpretation instead of better understanding.*

Storage

Excel is a splendid database for collecting and storing even relatively big amounts of data, and furthermore the program has a number of functions/functionalities for entering, sorting, filtering, extracting, retrieving and looking up data in the database. In the laboratory this is of course mostly relevant for data which are generated/accumulated repeatedly over time together with all the necessary information for further treatment and evaluation. *But what could at a first sight be looking like a nice and neat table for collecting data and information is not necessarily the optimal way of organizing data in Excel!*

Automation

Many employees in laboratories experience Excel as a kind of “black box”, where they type in data and just press “Calculate” before the program automatically takes the data to other parts of the workbook, does some calculations, presents the results, evaluates and comments these in relation to some preset conditions etc. etc. Somebody (maybe even years ago) has set up an application which does the job by using all the fancy functionalities within Excel but *the daily user may not understand what is really happening, how the procedure may be changed – and what if the conditions for the actual calculations and evaluation fail....?*



Presentation

Excel is not a great text editor but it has an almost endless toolbox for setting up text formats, orientation, backgrounds and colours etc. of the data in tables and charts before the result can finally be printed out as a nice report. For laboratories that have to deliver clear messages and information to be used as basis for important decisions *these many options for fancy layout must though be used with care to support and highlight the content of a report and not to steal the whole picture blurring the understanding of the real message.*

Excel IS a fantastic tool for data handling – also in the testing laboratory (although not developed especially for that purpose) – but it takes some deliberate policy in the laboratory to get the most out of it in an efficient, reliable and secure way.

E.g. by implementing the following principles/advice:

Planning – *a well spent effort before you start.*

Some claim that the "80/20 rule" is also valid for setting up a new spreadsheet application – i.e. use 80% of the time for planning by use of paper and pen and then afterwards the 20% for actually writing the application on the computer! Like...

- Which data has to be handled and how are they introduced in Excel (transfer, import, typing in)?
- Which calculation needs to be done, what are the formulas and which inputs are needed (incl. any constants)?
- How must the result be presented and are there any special requirements with regard to QA and documentation?
- Is the application going to be used routinely – or is it just a single event – and who is going to handle it?
- If the purpose is storing and organizing data a suitable table ("list") must be set up before entering the data..... a.o.

Setting up applications – *to be used by more than "one nerd" in the laboratory.*

Think of the usefulness of an Excel application by e.g. ...

- Making it clear where data-input must be done and how
- Making the use of an Excel application self-explanatory (on the level where the daily user needs to understand the actual functionality of the spreadsheet)
- Use named cells and cell ranges which makes formulas easier to read and understand afterwards
- Hide (if possible) some background calculations in hidden cell areas of specific "calculation sheets"
- Build in automatism (by using logic functions, conditional formatting, data validation – and if relevant even array functions and macros) where it doesn't conflict with the users' logical judgements and need to act individually
- Keeping an overview of big amounts of data by splitting them up in different sheets and files – and linking these in a convenient way (... which at the same time will possibly reduce time consumption for calculations)
- Securing (protecting) cells/ranges/sheets before storing the application as a template for future use a.o.

Documentation – *the irritating, time consuming but invaluable "paper work"*

Besides being a required part of any QA system under which the Excel applications are used it can prove being of crucial importance for the daily use that the functionality of a given application is well-documented. Not least when the colleague who developed the various formulas and other functionalities is not available anymore. Excel has a number of more or less elegant features for this

- Write explanatory text in cells adjacent to the formulas (use e.g. the equation editor to show real formulas)
- Use the comment tool to write explanatory text to the content of cells
- Make a print out of the worksheet in "Show Formulas" mode...a.o.

Quality Assurance – *also spreadsheets must be brought to follow the rules*

As an integrated – and very important – part of the whole testing process in the laboratory the calculation tool in form of Excel must of course also be a means of quality assurance activities, like validation, documentation etc.

- Test the functionality of an application by manual check calculations
- Use (if necessary) the built-in facilities for step by step calculations and error-tracking
- Audit trails of corrections in a spreadsheet are unfortunately not included in the functionality of Excel, but must be established by creating special macros.

[A part of the QA around the use of Excel will of course also be formulating and following a policy and some written procedures; see the following]

Policy/procedures – *the laboratory must be in full control of the use of this powerful tool.*

Many employees may find working with Excel and developing new and better (!?) applications an interesting – and even fun – challenge, which of course can be seen as positive but at the same time can create a lot of problems with regard to controlling the actual and consistent use of Excel in the laboratory!

For all uses of Excel for data handling in a testing laboratory it is therefore crucial that the laboratory has set up its own policy and relevant procedures to monitor (limit) the temptation to develop applications in all kinds of fancy directions. This must/could e.g. cover subjects like...

- Who is responsible for developing (and revising) the applications used for data handling in the laboratory?
- For which purposes is Excel used in the laboratory?
- How (on which level) are Excel workbooks secured?
- How are statistics dealt with in the laboratory?.... a.o.

Training – *you can always learn something more about Excel*

Excel is often seen as a tool which people can easily adopt and learn by themselves! It is though often advantageous for the whole organisation in a laboratory if the staff gets specific training on the use of Excel for laboratory applications (many of the specific functionalities, functions and ways of developing a spreadsheet application relevant for laboratories are not taught in normal Excel courses).