

InnoRenew CoE

Livade 6, 6310 Izola/Isola, Slovenia, T: +386 40 282 944, E: coe@innorenew.eu, www.innorenew.eu

Best practices for using spreadsheets in research data management

Dr. Ana Slavec, InnoRenew CoE

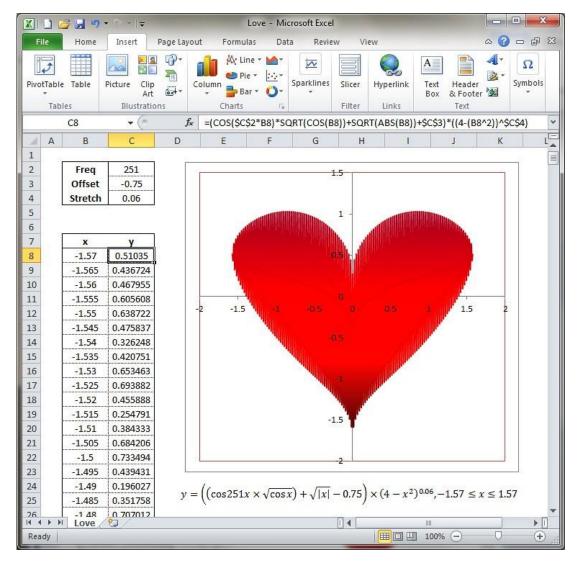
One in Five Genetics Papers Have Errors Should we use Excel spreadsheets? **Caused by Microsoft Excel** The Economist Spreadsheet programs like Excel are creating issues when they reformat scientific data. spreadsheet Risks in Science By Avery Thompson Aug 26, 2016 Excel errors and science papers Neuroskeptic | By Neuroskeptic | August 6, 2017 5;08 PM Spreadsheets are playing havoc with scientists Graphic detail > Sep 7th 2016 | by THE DATA TEAM 00000 f 🗸 🥳 🖸 🖾 THREE years ago Thomas Herndon, a young graduate student from the University of Massachusetts Amherst, made a splash among economists. Given an assignment to repli-Published academic the use of spreadsheets such as Microsoft Excel could pose risks for science. That's g to a preprint posted on arXiv from Ghada AlTarawneh and Simon Thorne of Cardiff plitan University. AlTarawneh and Thorne conducted a survey of 17 researchers from the Excel spreadsheets might be sity of Newcastle neuroscience research centre, ranging from PhD students to senior chers. None of the respondants had any formal, certified training in spreadsheet use, with ruining science for everyone According to a study in the journal Genome Biology, around 20 percent of genetics (71%) being self-taught. Despite this, the majority of the researchers rated themselves as have errors that have been introduced by Microsoft Excel or similar mediate" users of Excel. Worryingly, only 20% reported that they had their spreadsheets The researchers analyzed more than 35,000 genetics papers An alarming number of scientific papers contain Excel errors ked for errors by colleagues ('peer reviewed'). Most said they did all the testing themselve One more reason to hate Excel. Tibi Puiu - August 7, 2017 - Updated on August 8, 2017 in News Submit to Reddit Share on Twitter ven that they did none at all. How did you learn how to use spreadsheets? Self-tutored Start your day by asking for Peer learning Certified training courses The Washington Post Self-tutored [12] August 26, 2016 at 12:17 p.m. GMT+2 A surprisingly high number of scientific papers in the Sold



Yes, because...

- It's intuitive, easy to use and flexible
- It's very popular

 Errors are usually due to misuse



Source: https://www.flickr.com/photos/7438170@N06/6862995461/



...but be aware that:

- XLS is a proprietary format
- Excel is error-prone
- Excel is not reproducibility-friendly





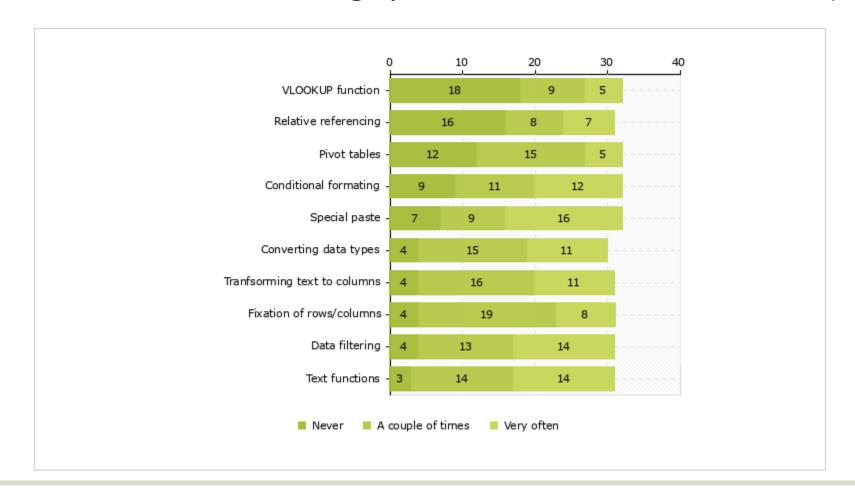
10 Excel functionalities I most often use

- Transforming text to columns
- Fixation of rows/columns
- Converting data types
- Data filtering
- Relative referencing

- Special paste
- Conditional formating
- Text functions
- Vlookup function
- Pivot tables



Have you ever used the following spreadsheets functionalities? (n=30)

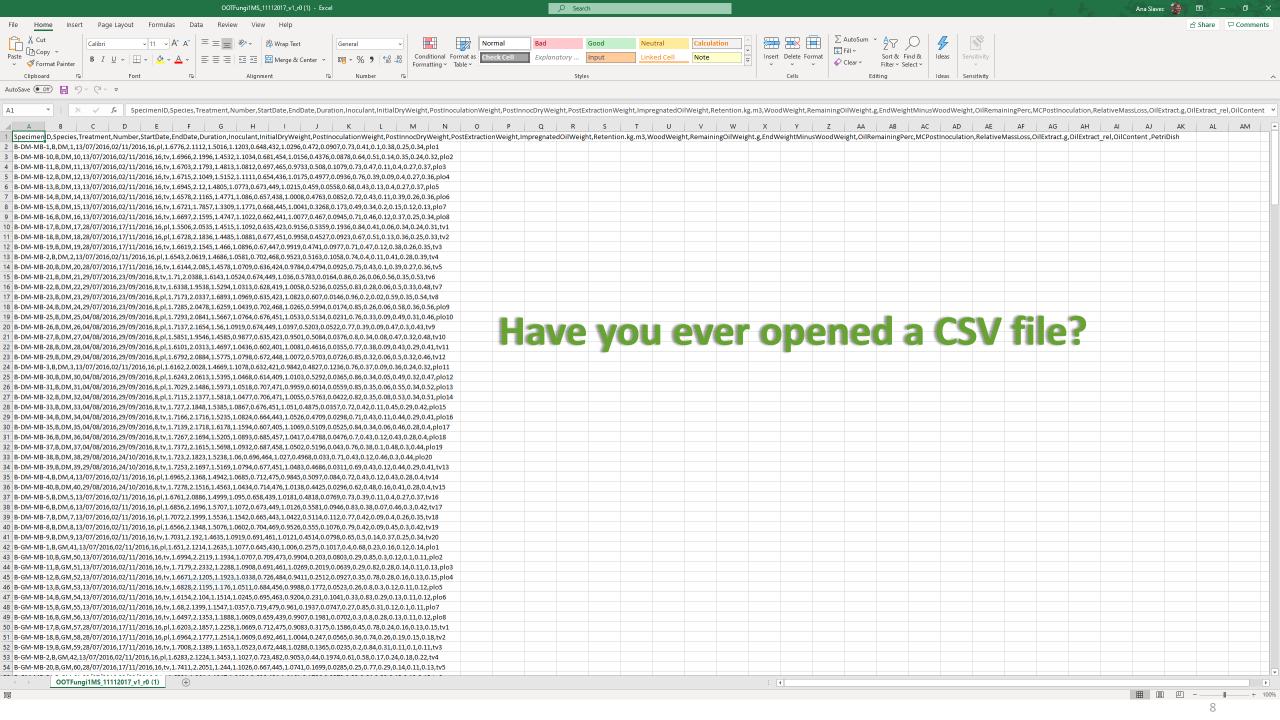


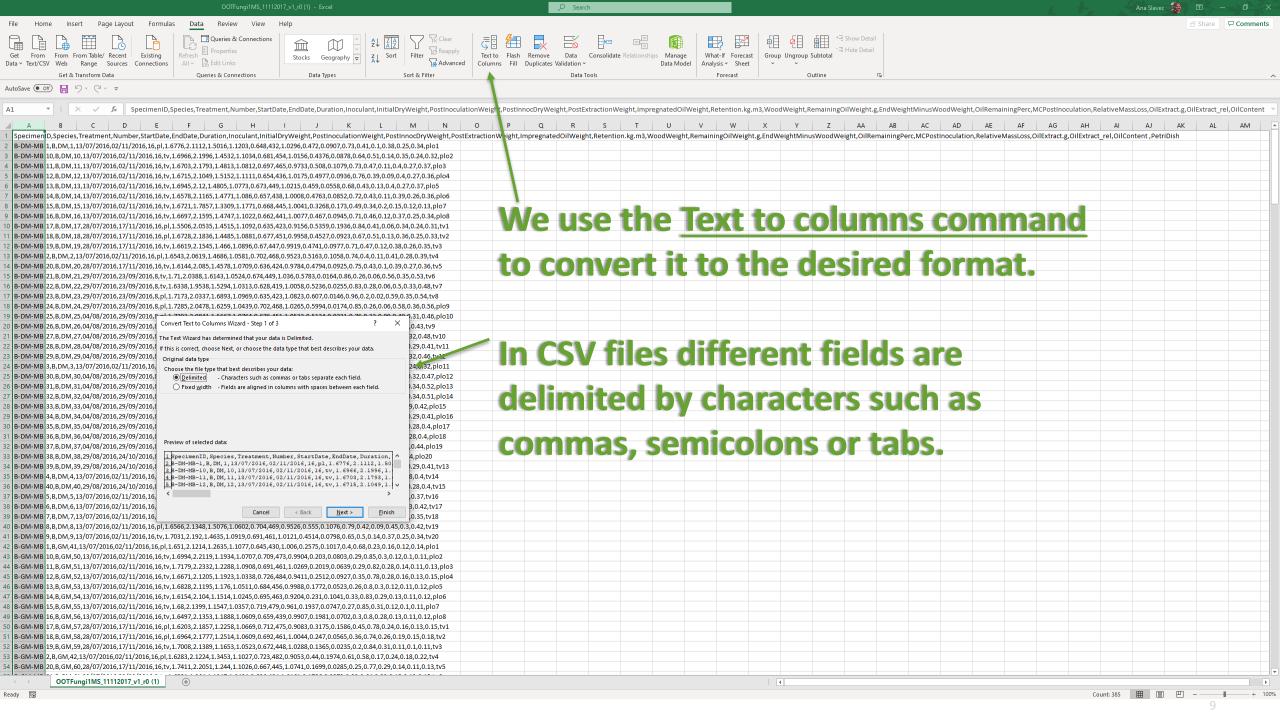


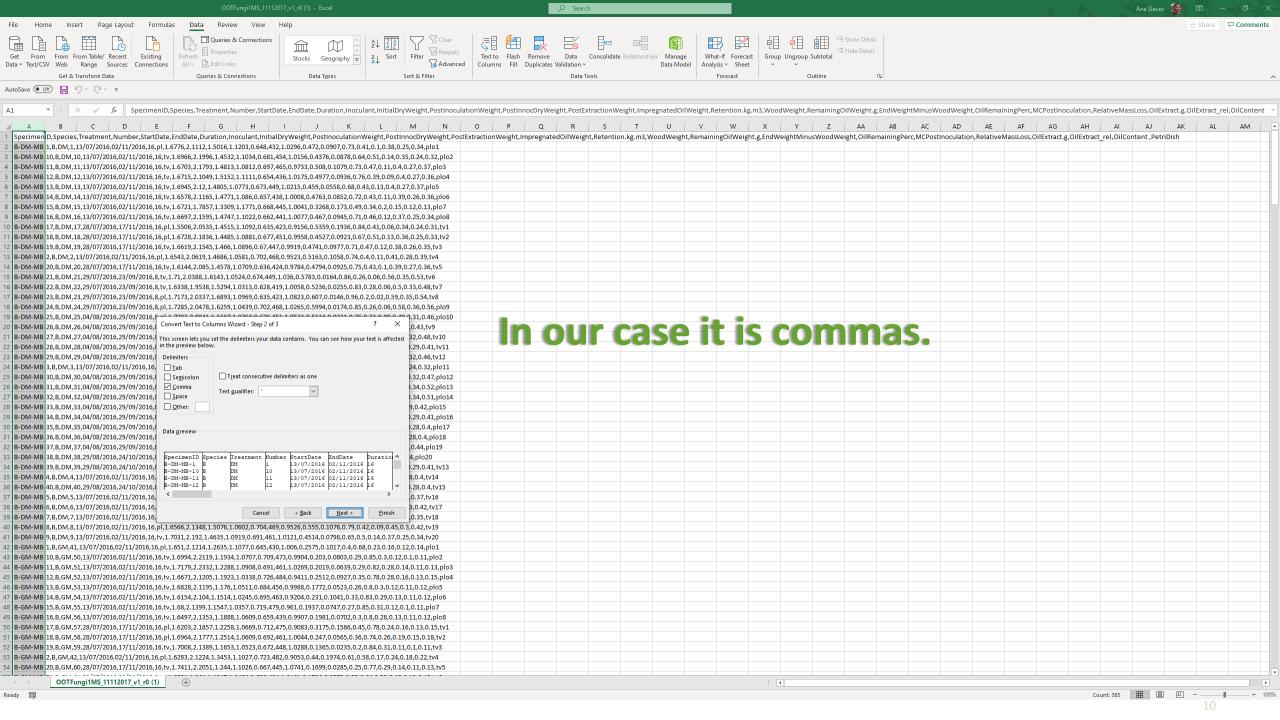
Dataset for exercises

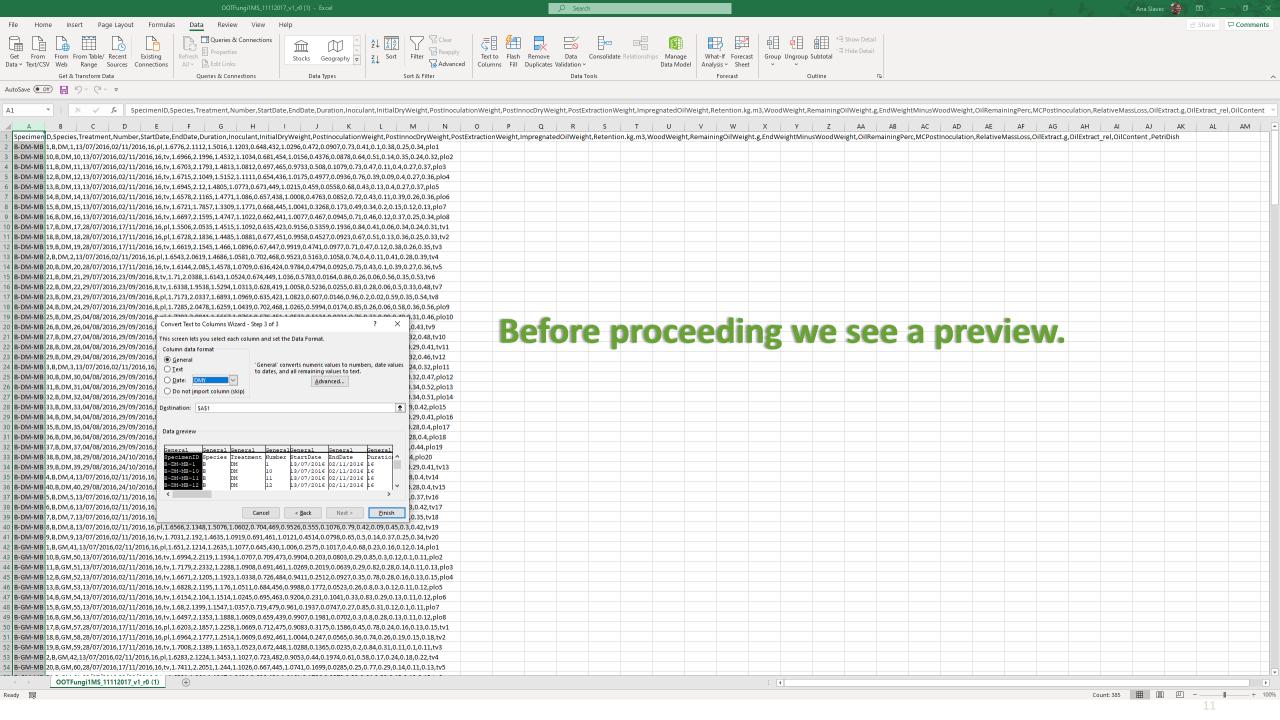
File: OOTFungi1MS 11112017 v1 r0.csv

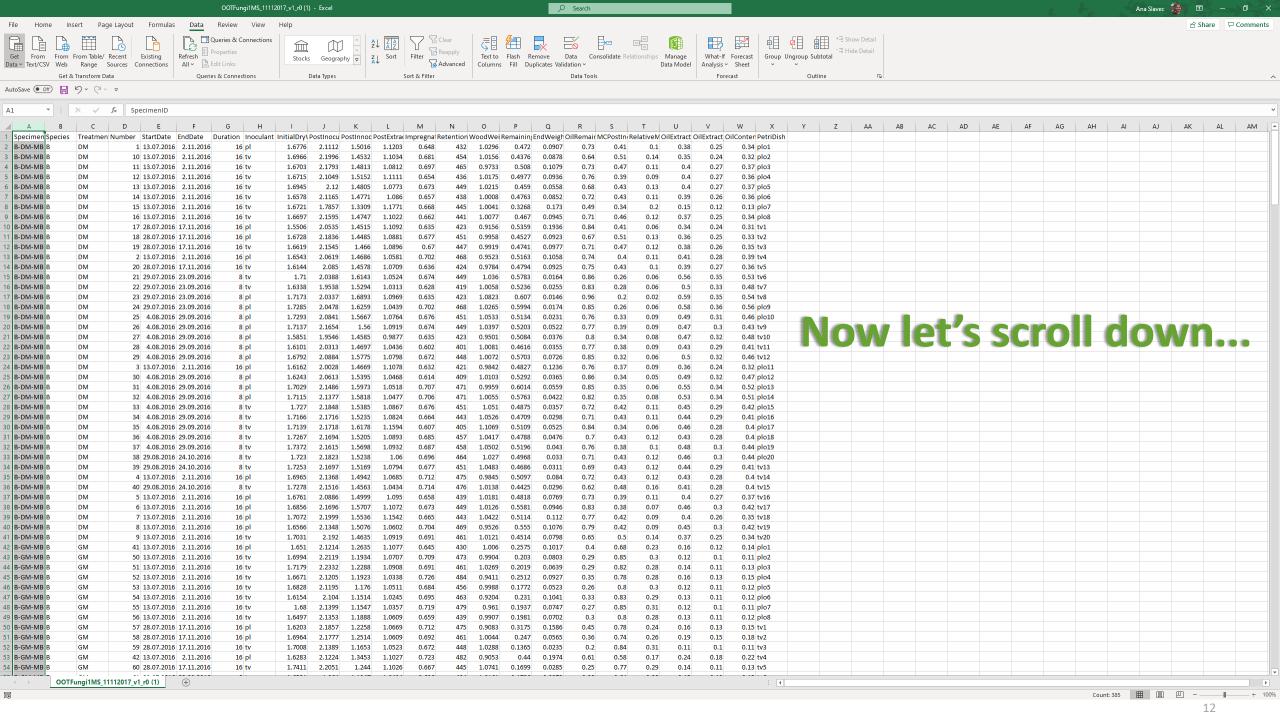
Source: Burnard, Michael, Schwarzkopf, Matthew, Tverezovskiy, Viacheslav, Treu, Andreas, Humar, Miha, & Kutnar, Andreja. (2018). Wood preservatives utilizing low-value olive oil production by-products: Analysis [Data set]. Zenodo. http://doi.org/10.5281/zenodo.1248463

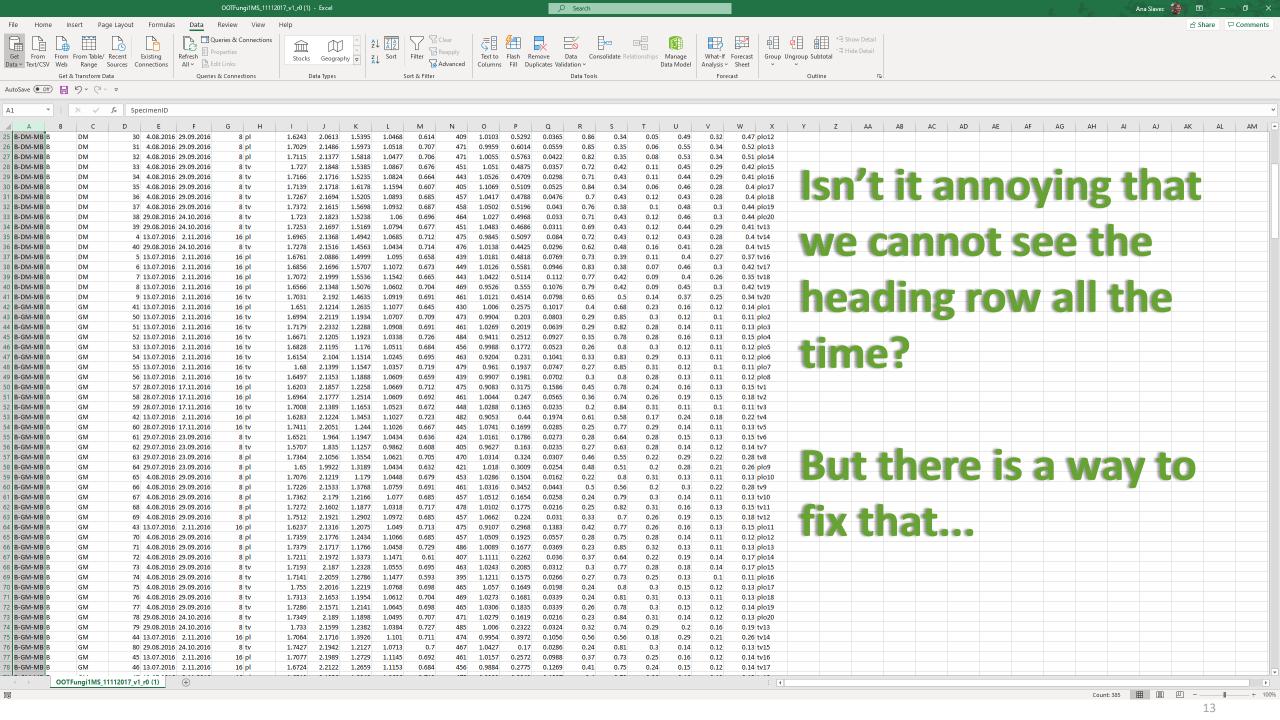


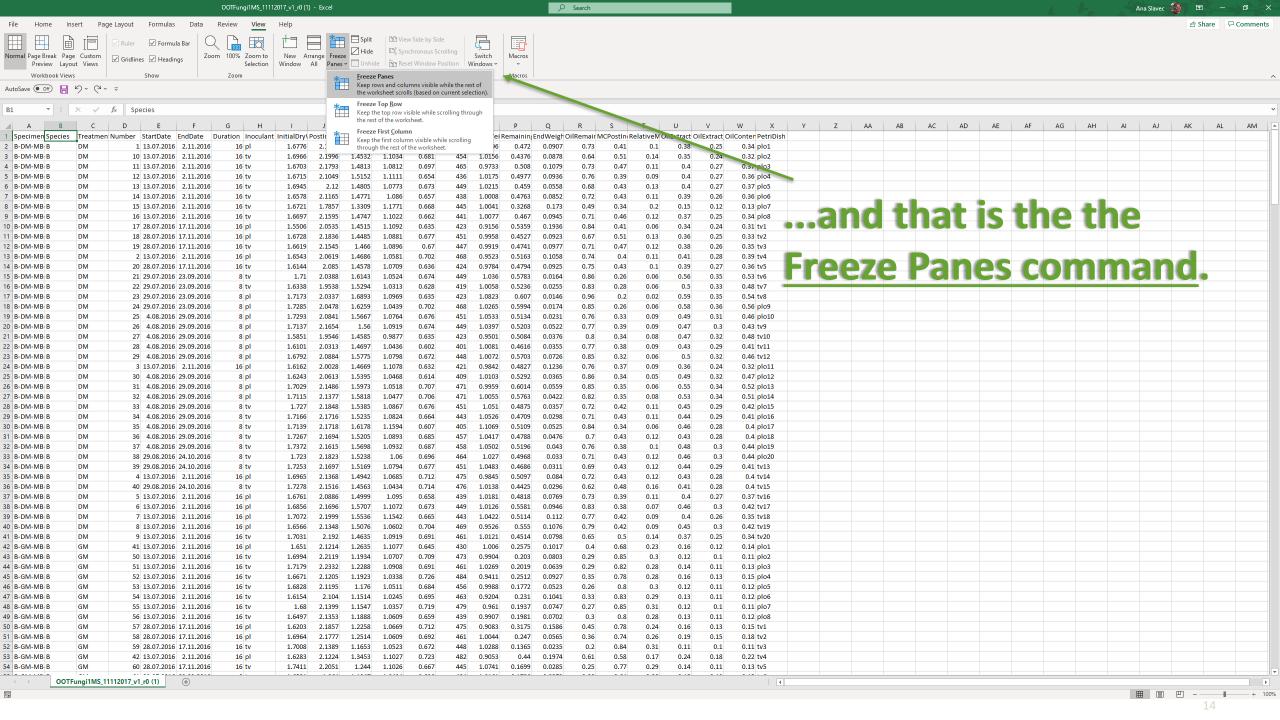


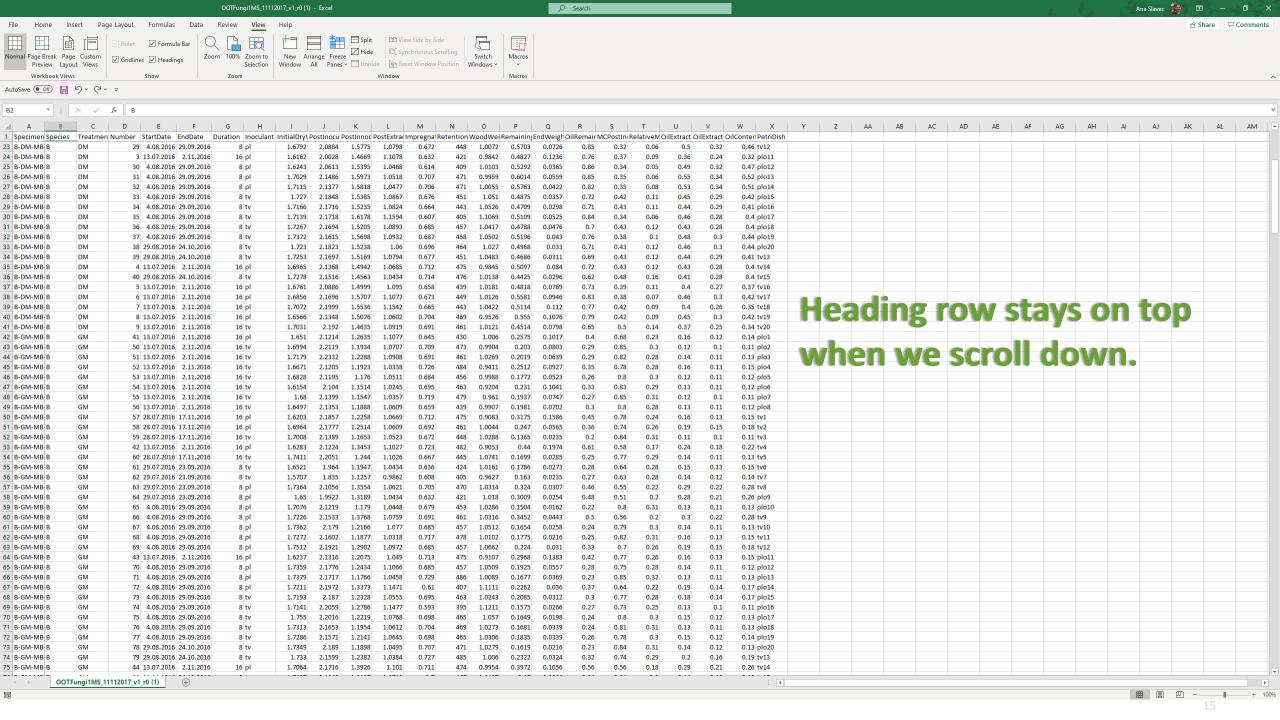


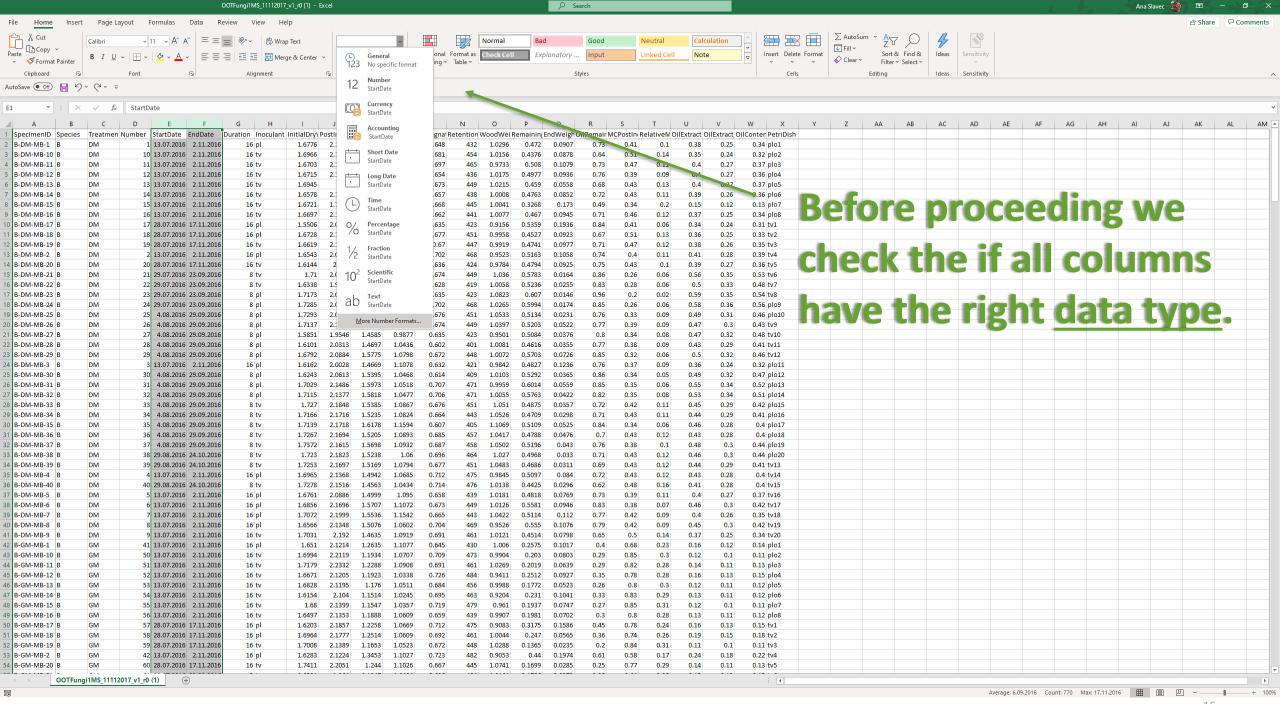


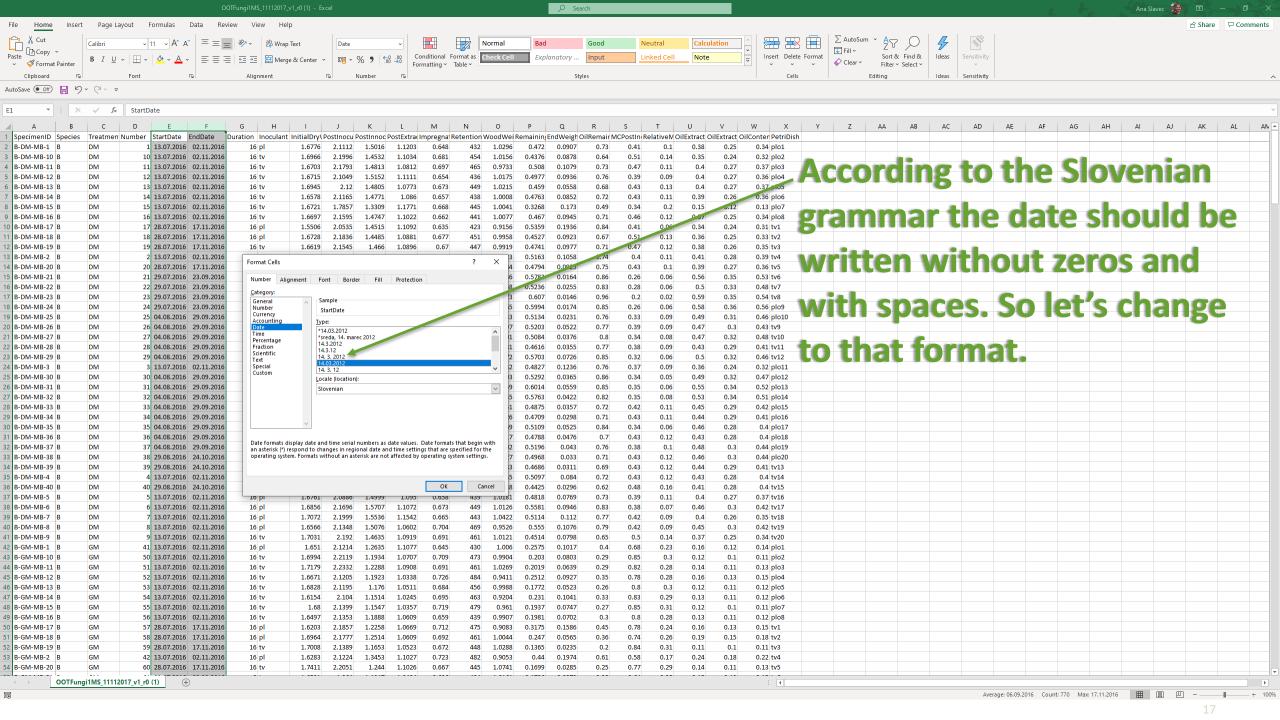


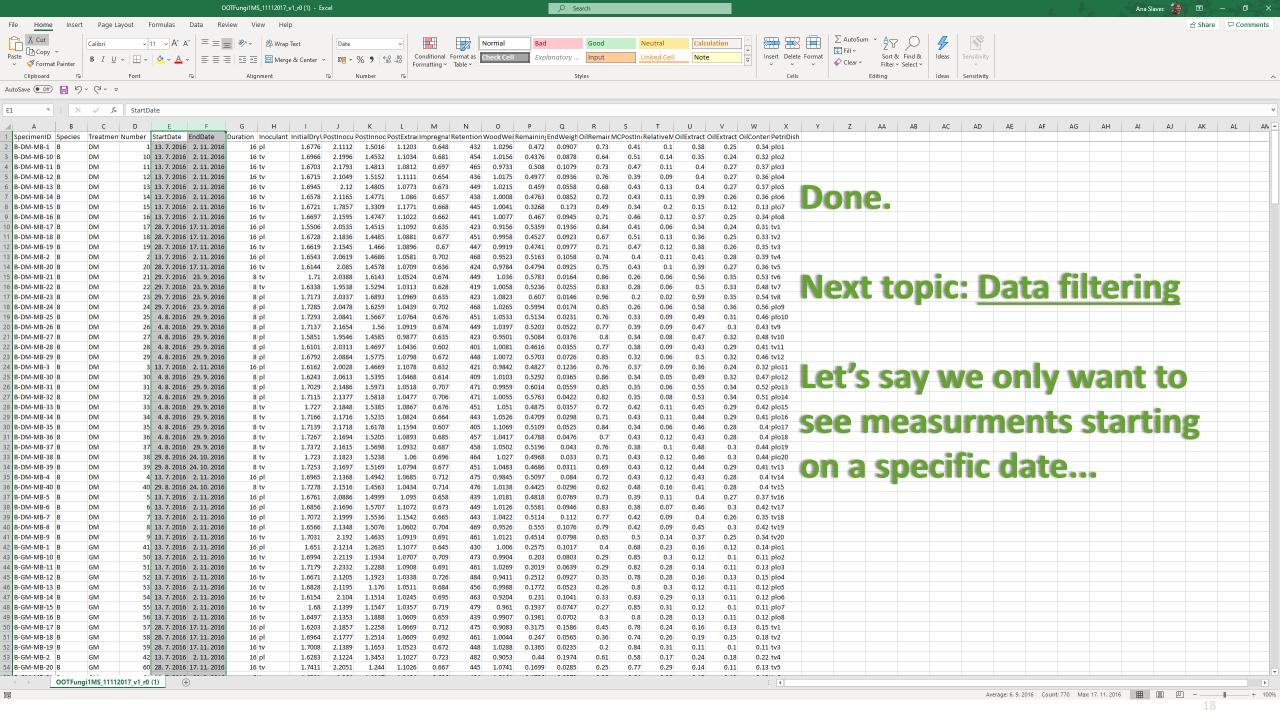


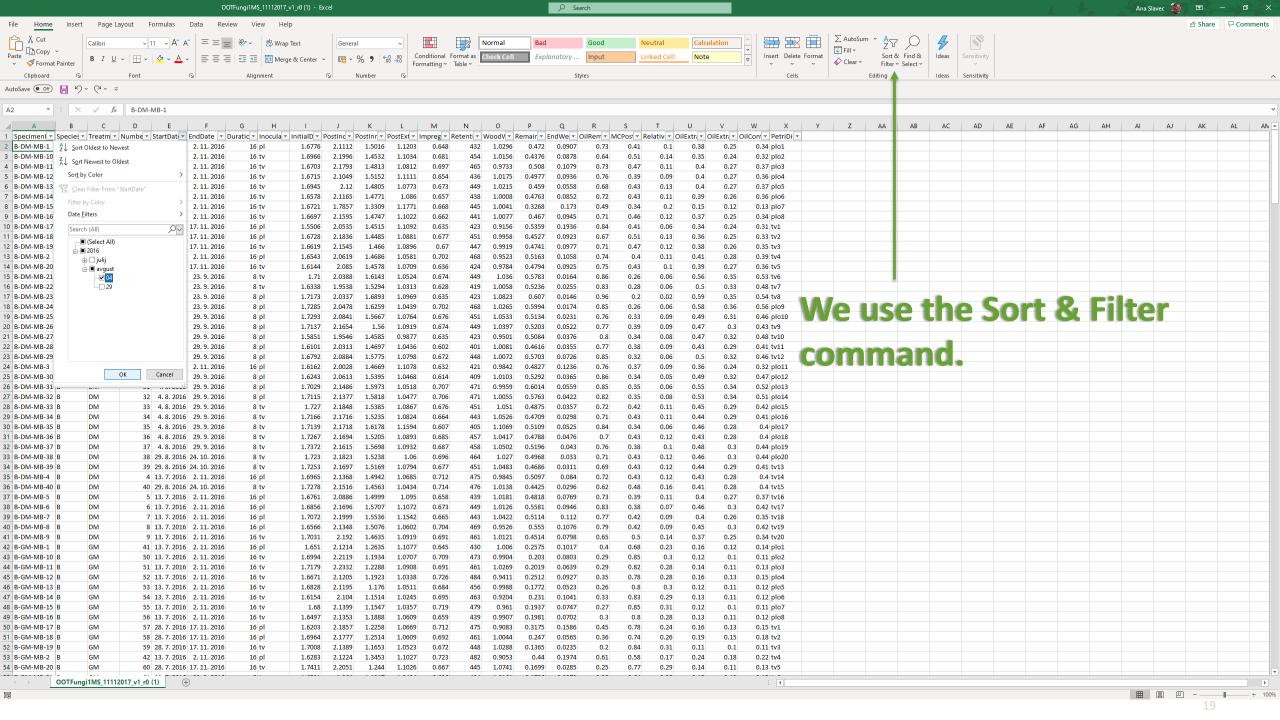


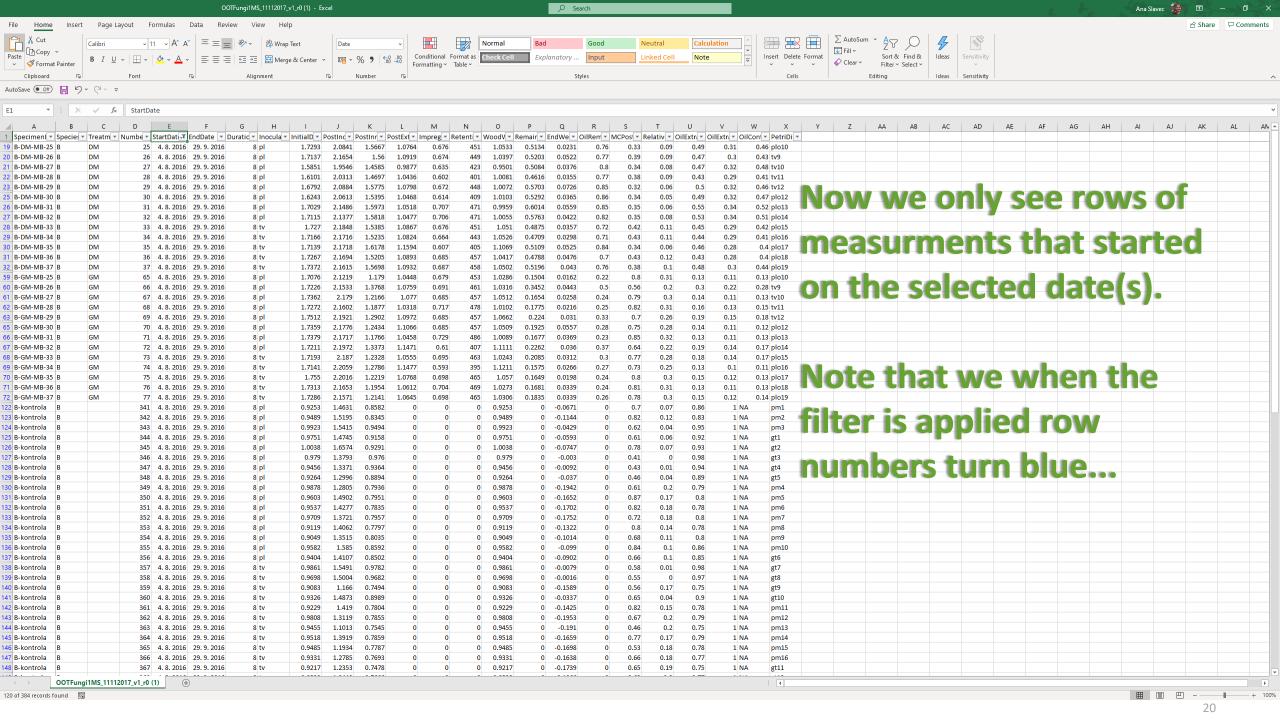


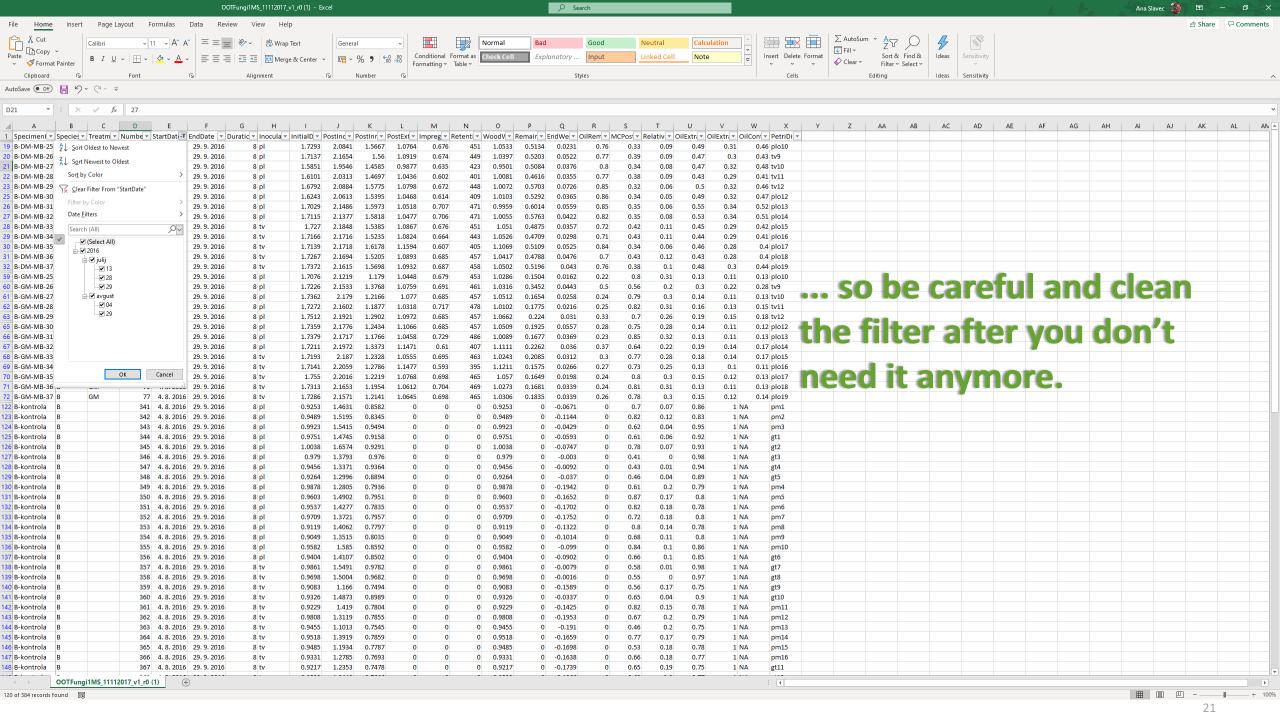


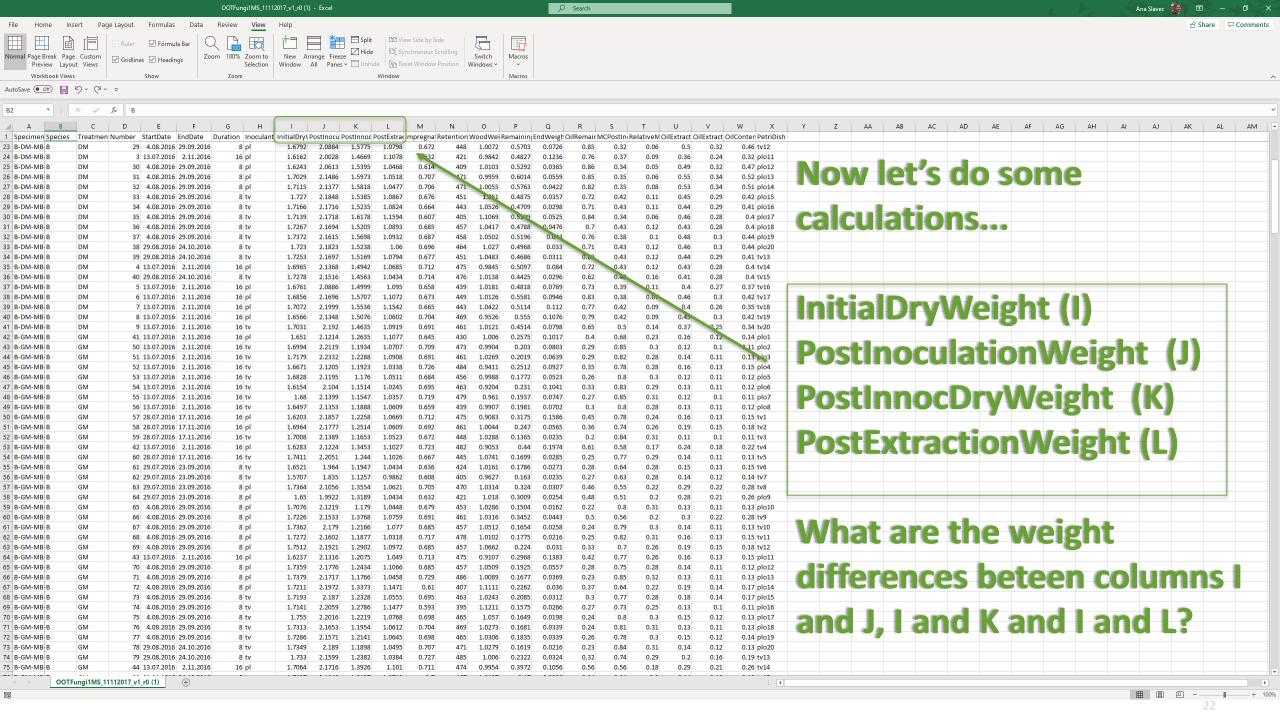


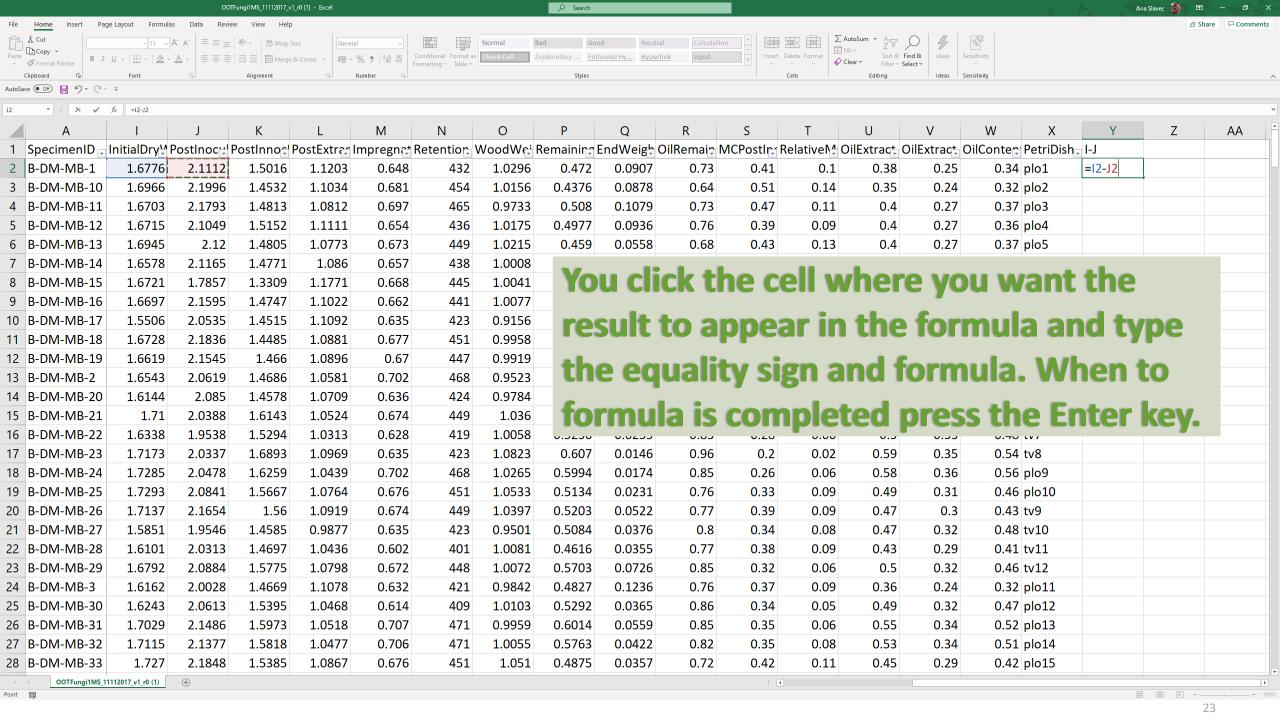


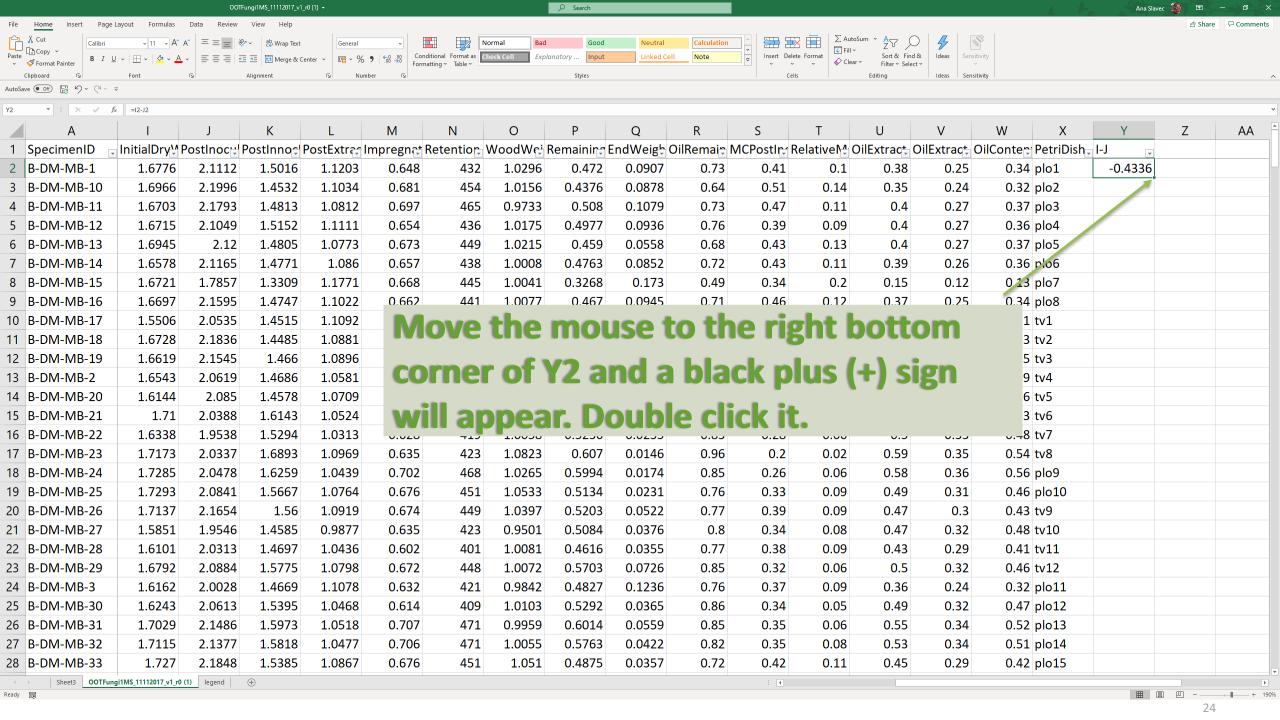


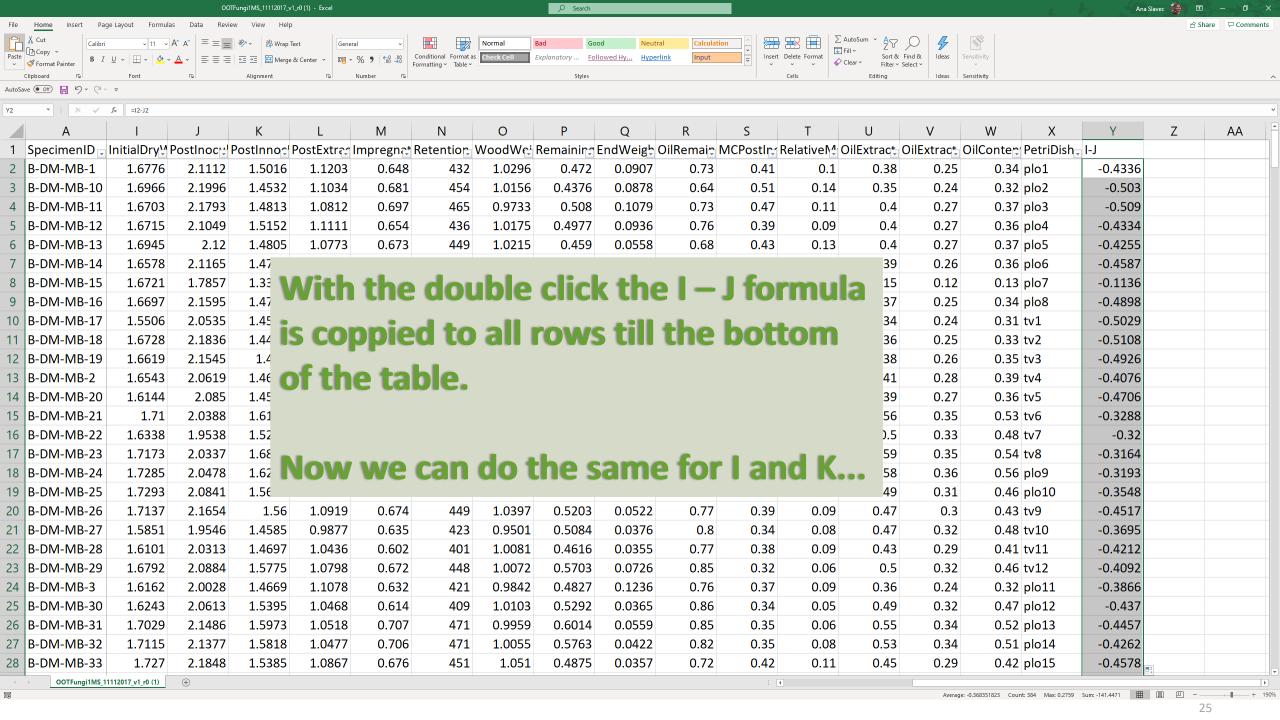


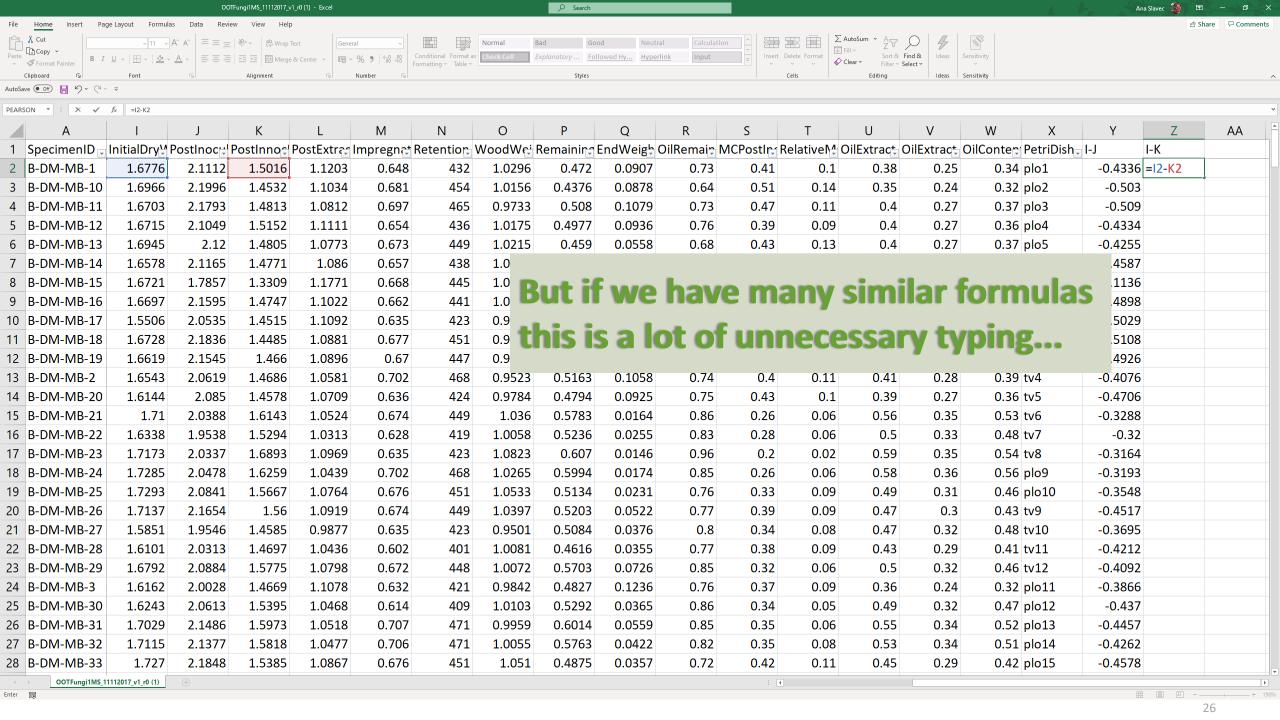


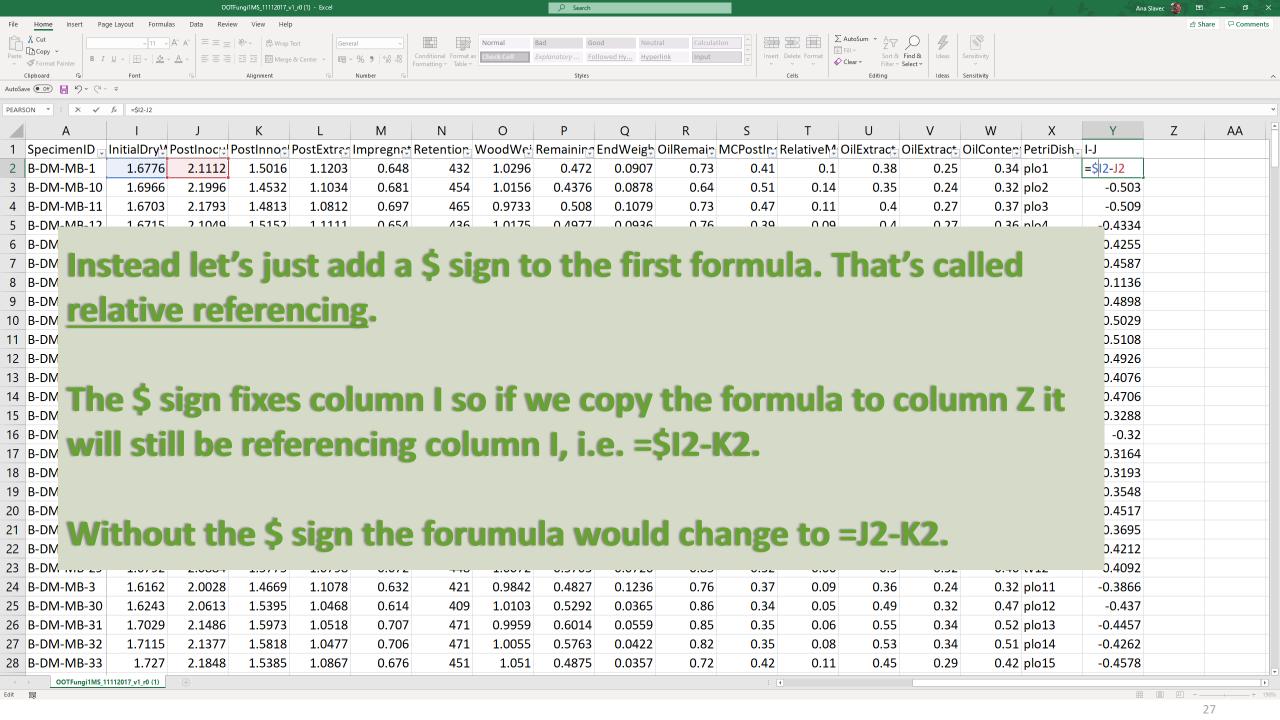


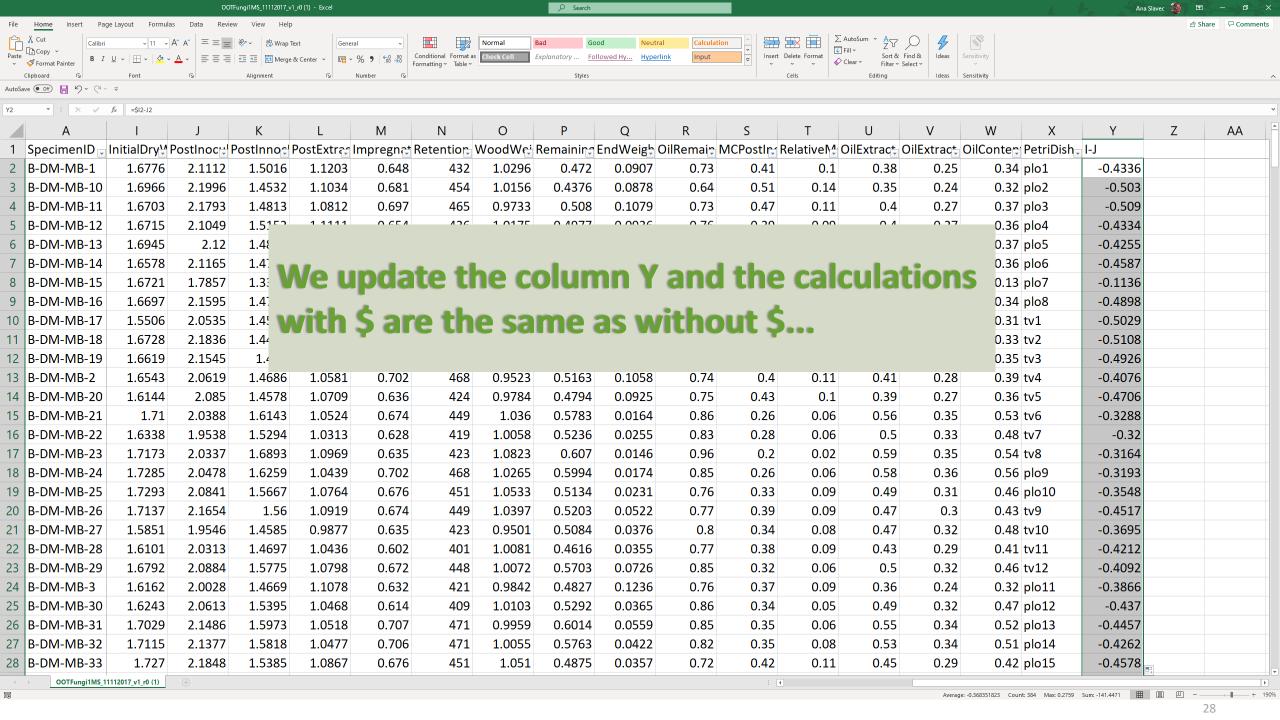


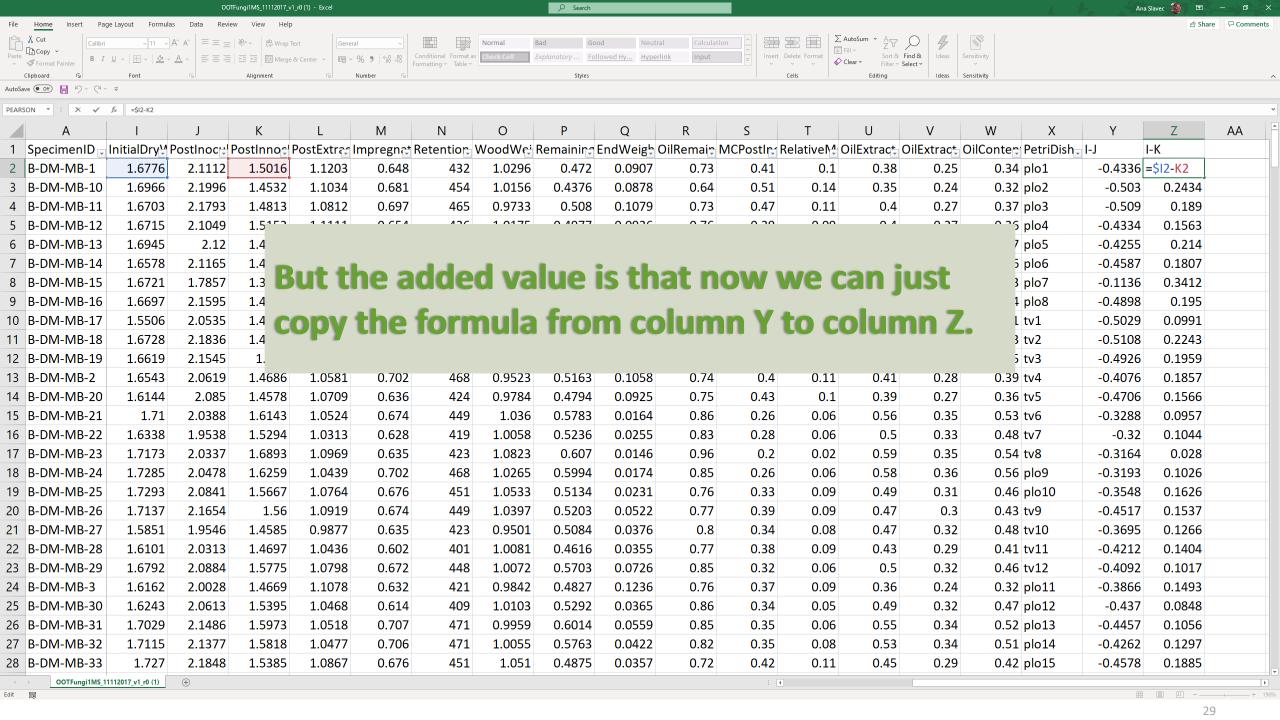


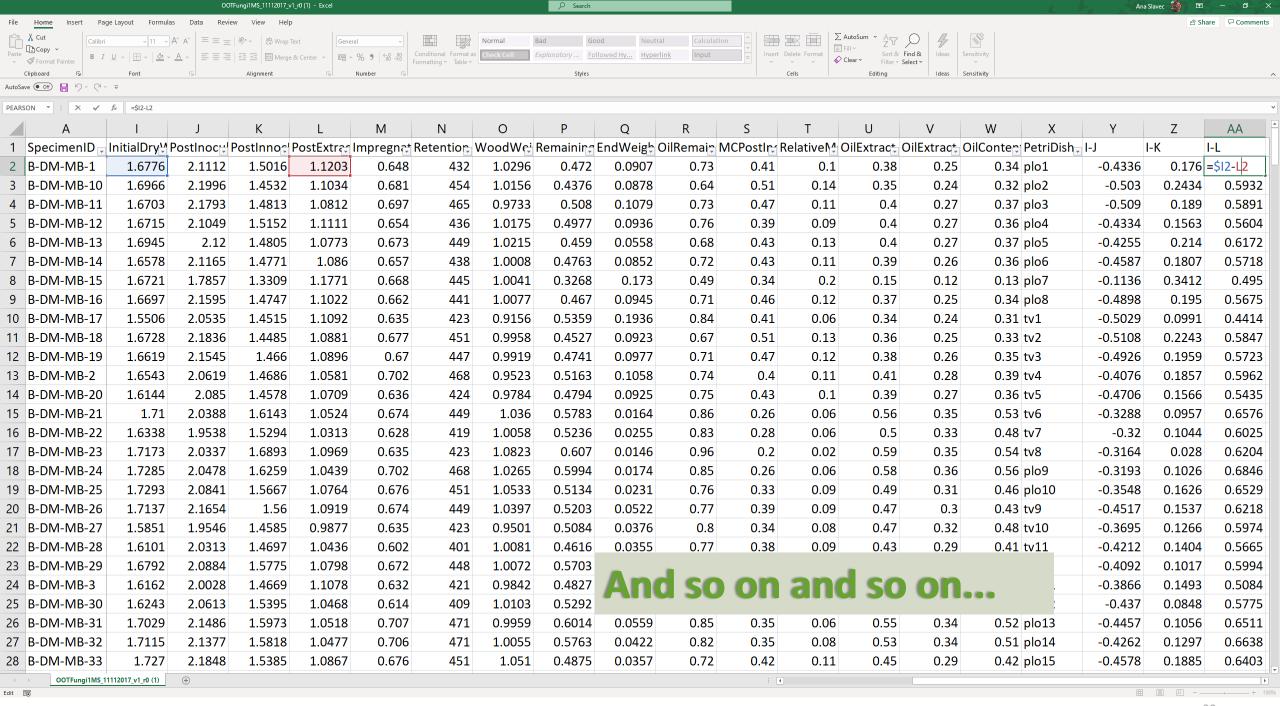


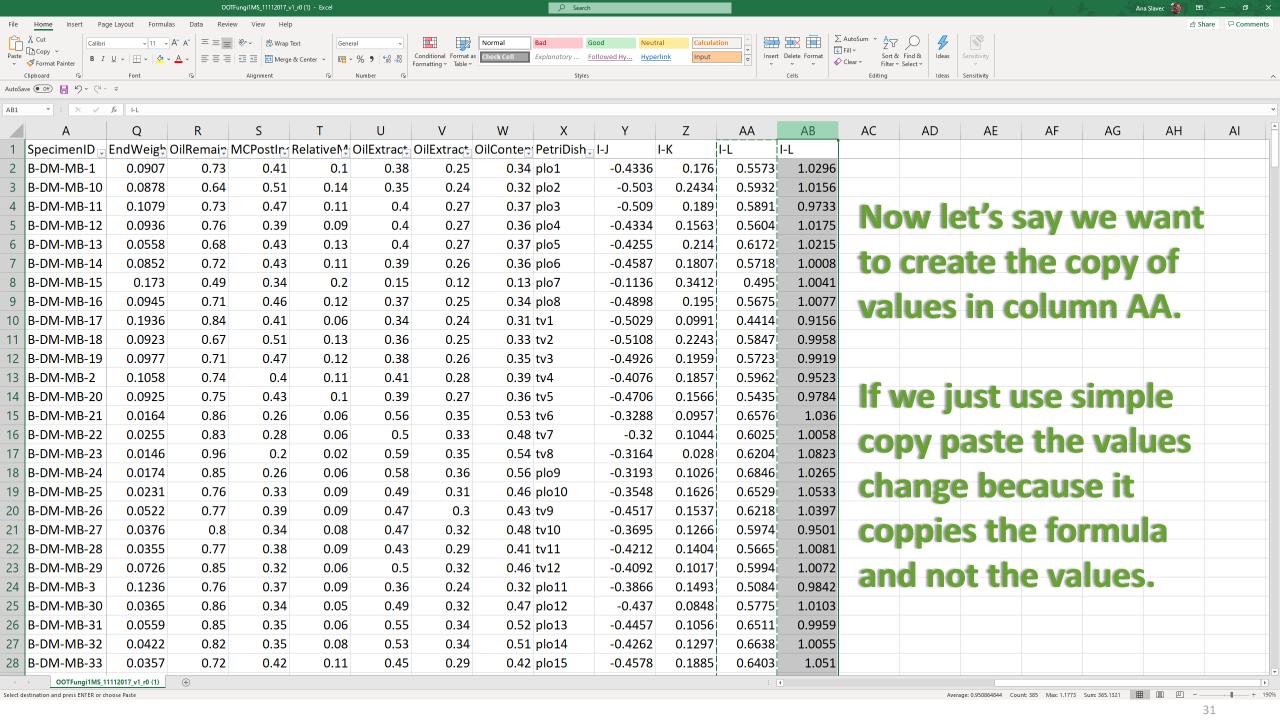


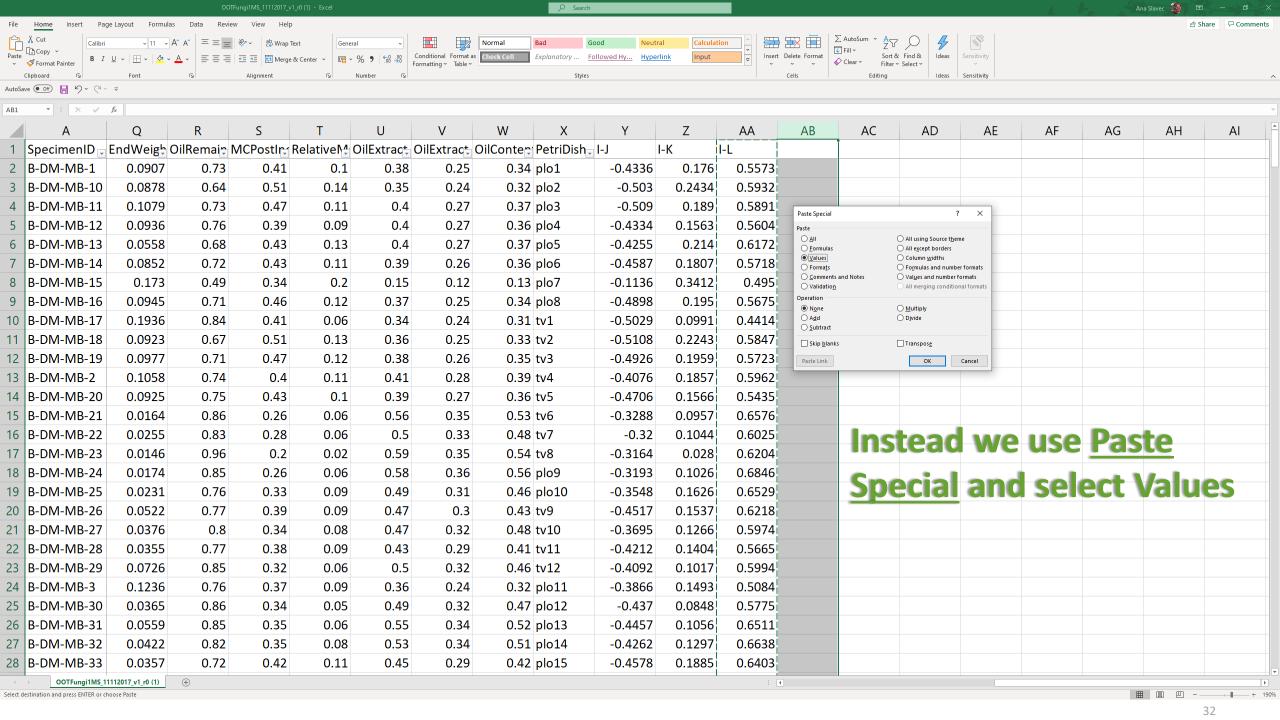


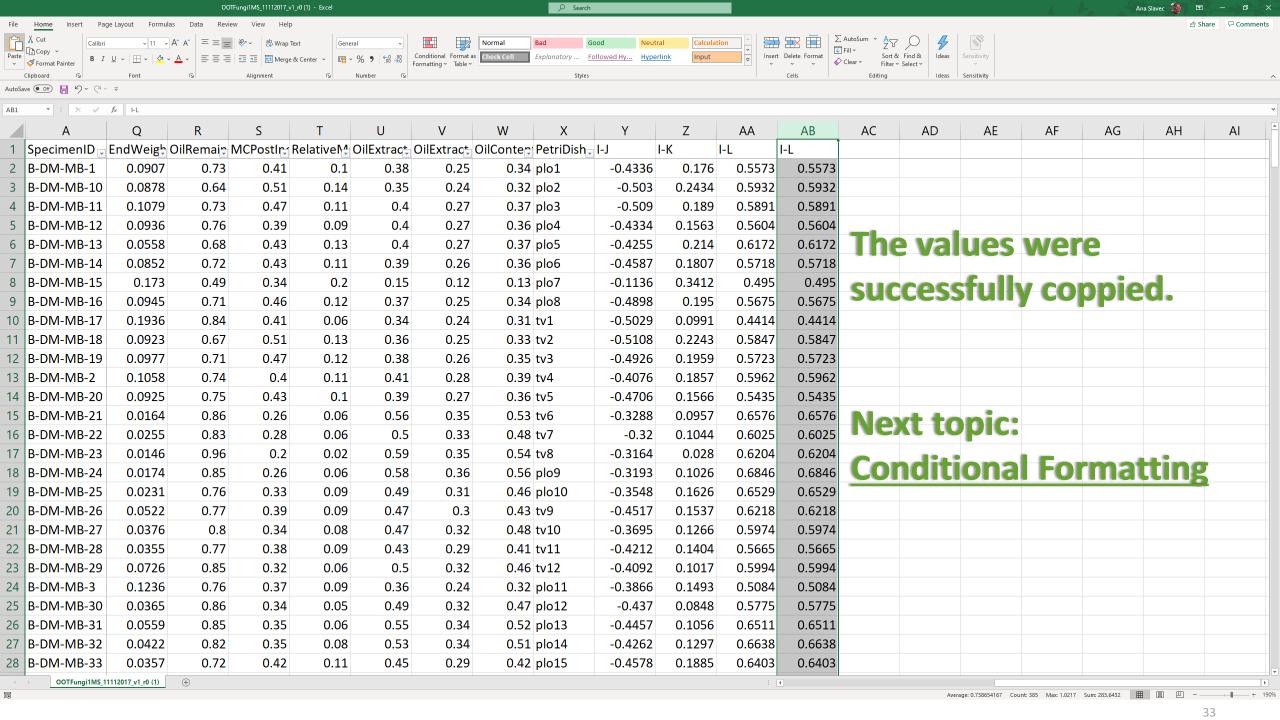


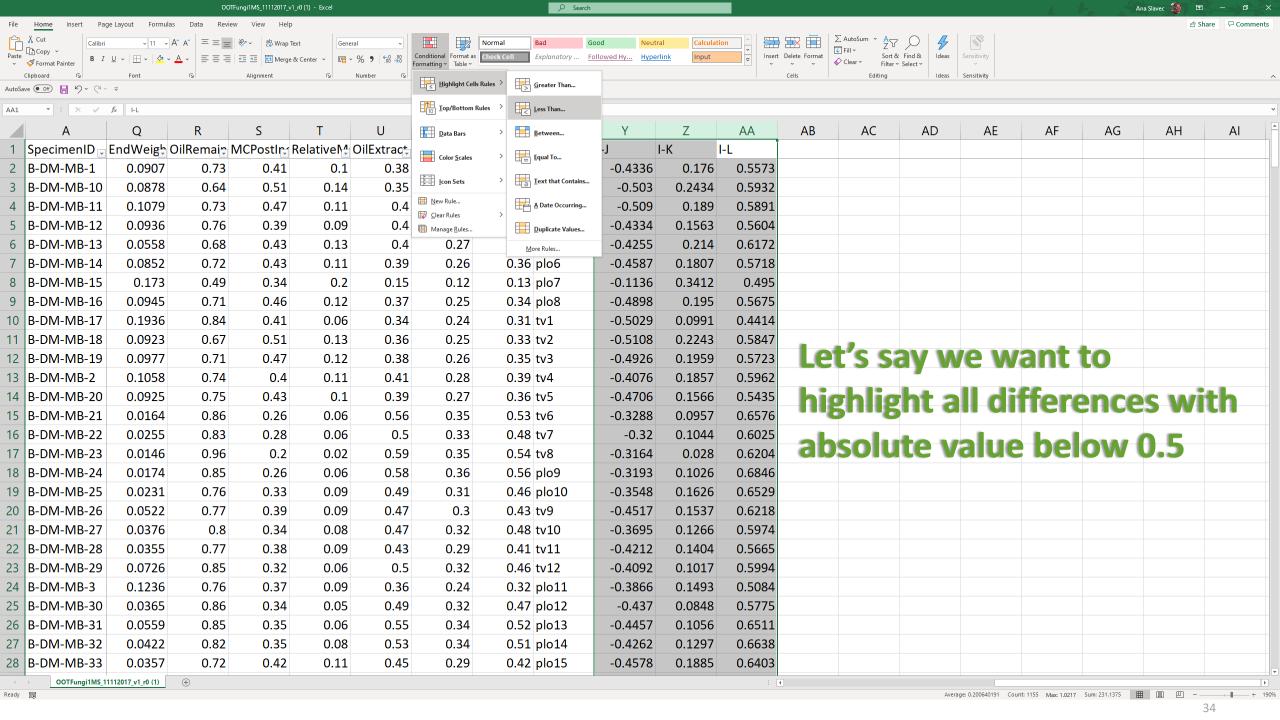


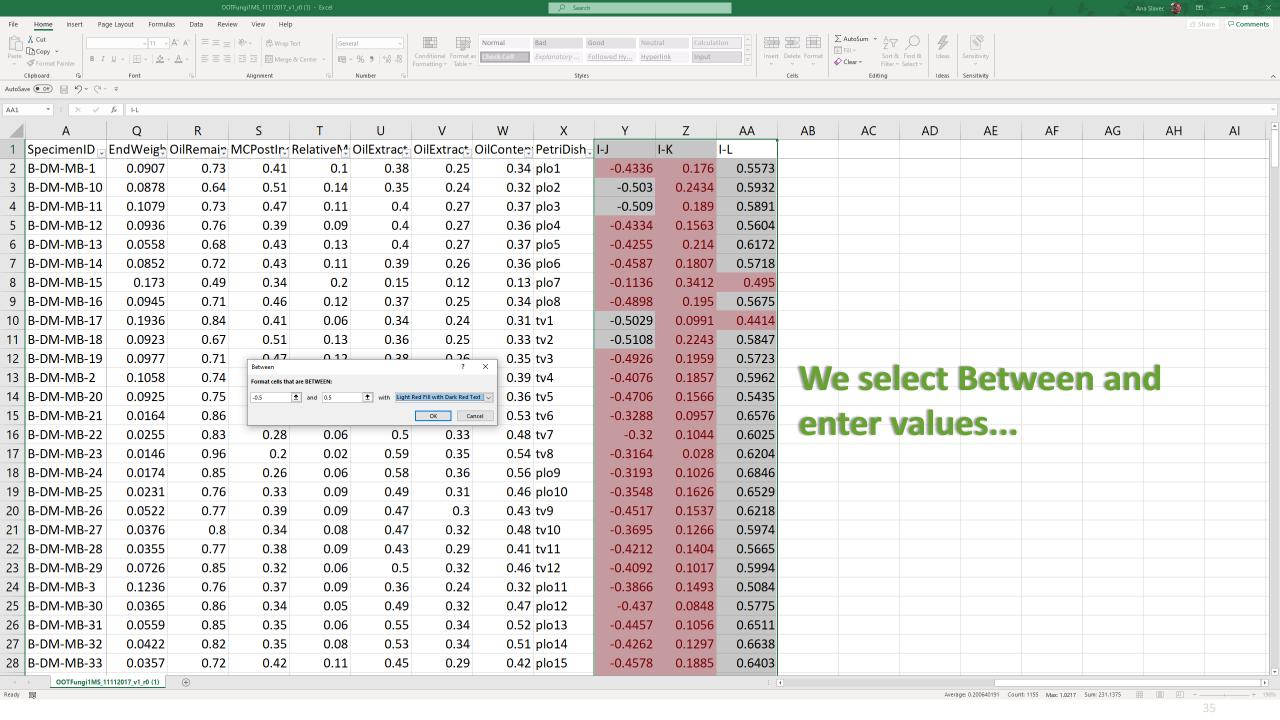


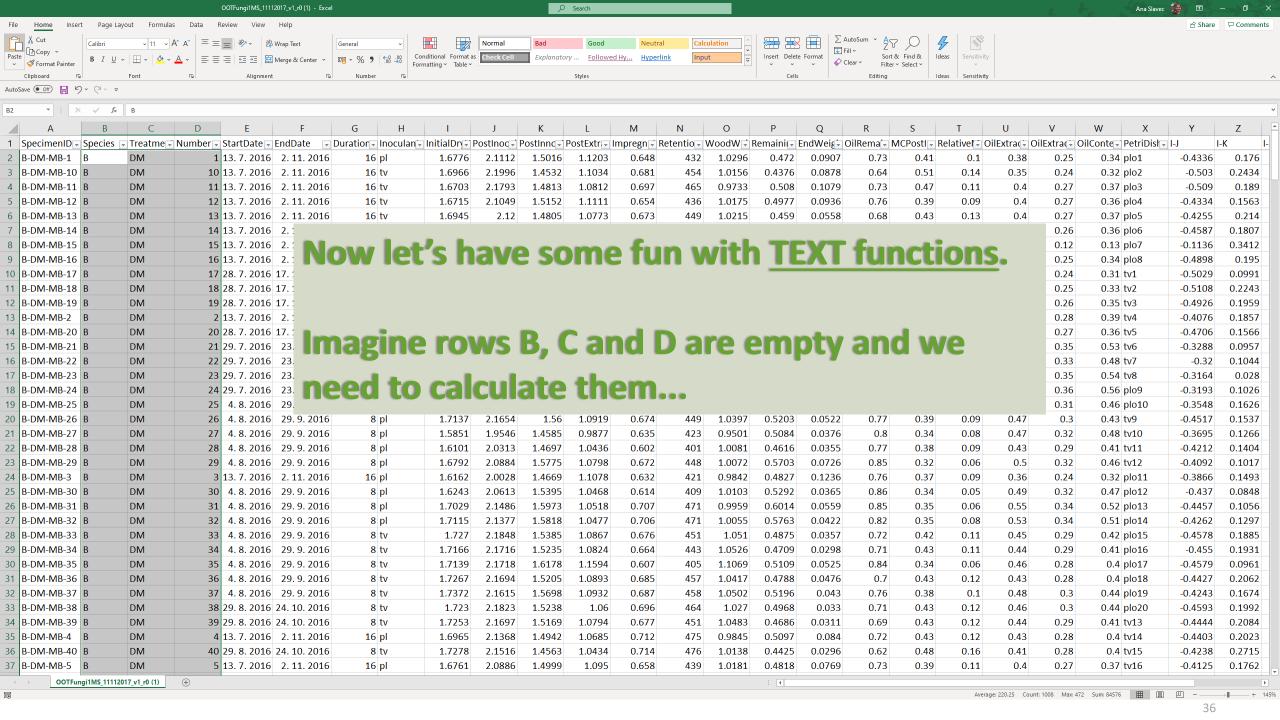


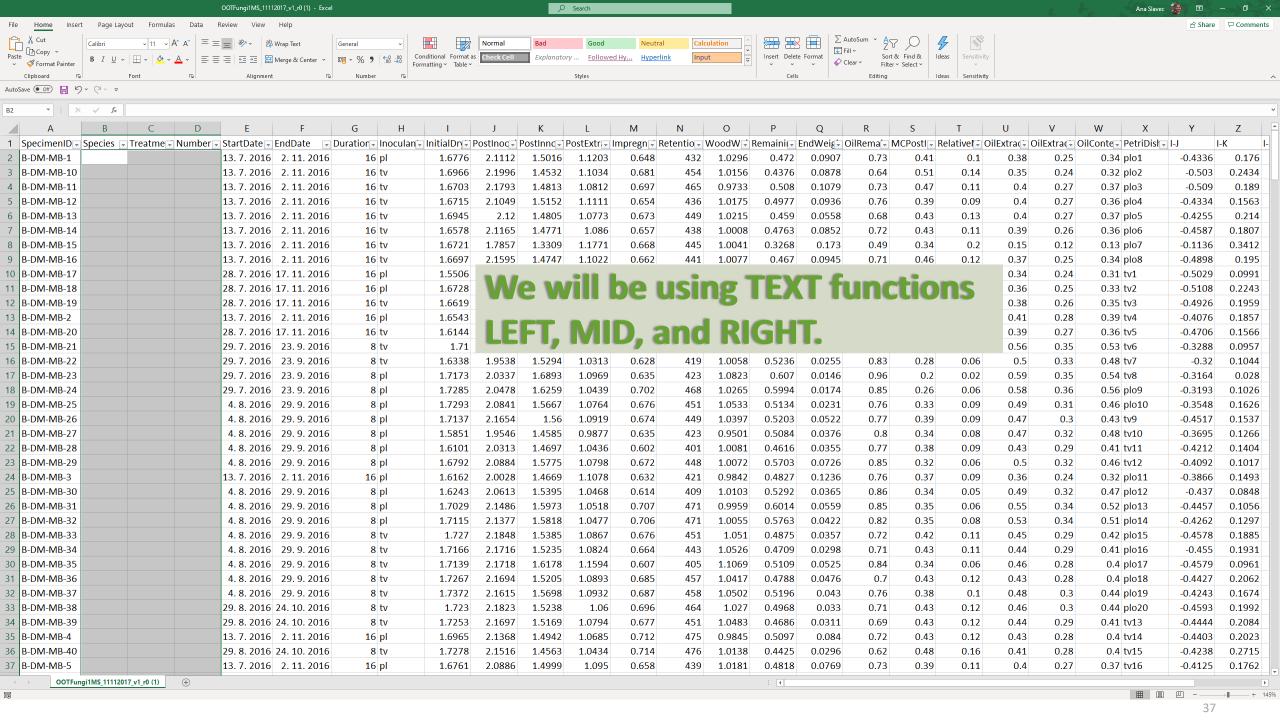


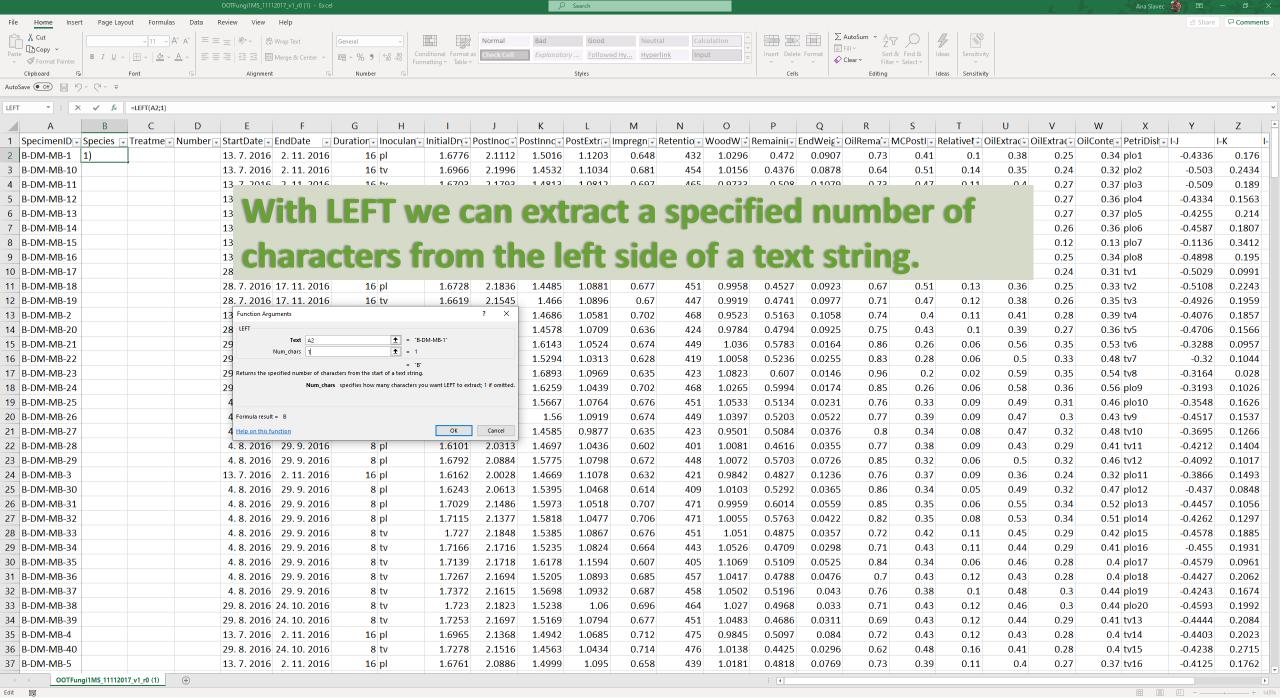


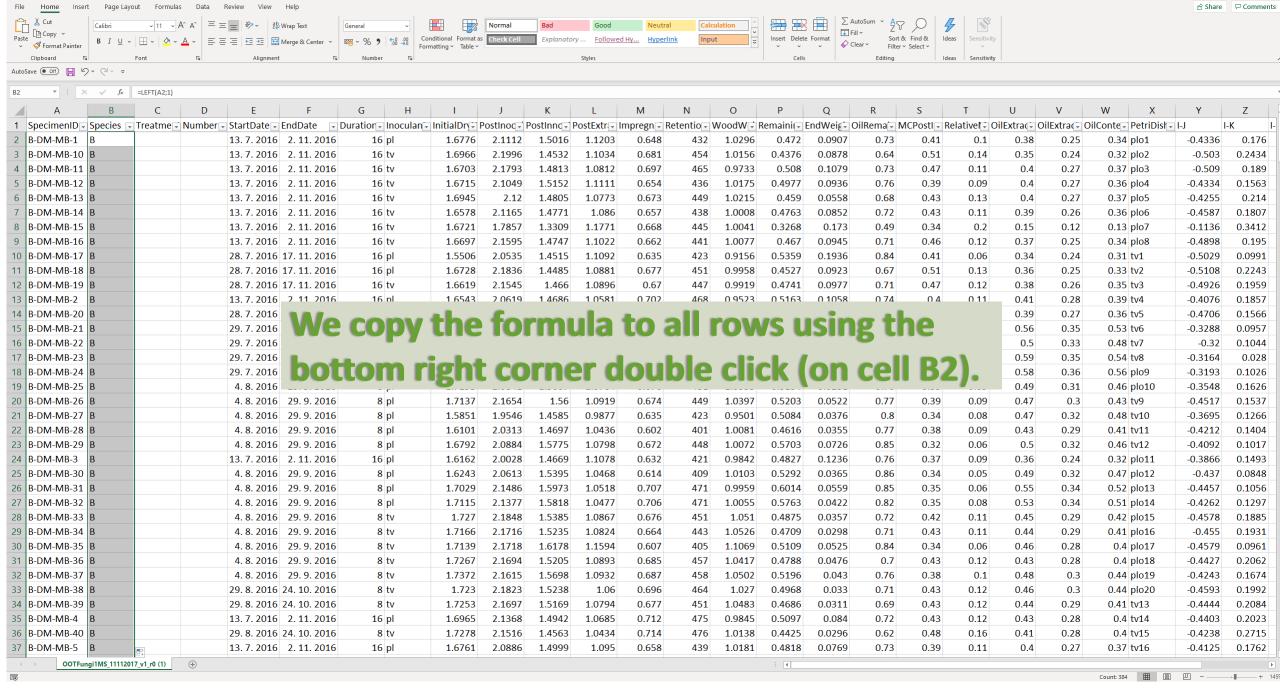


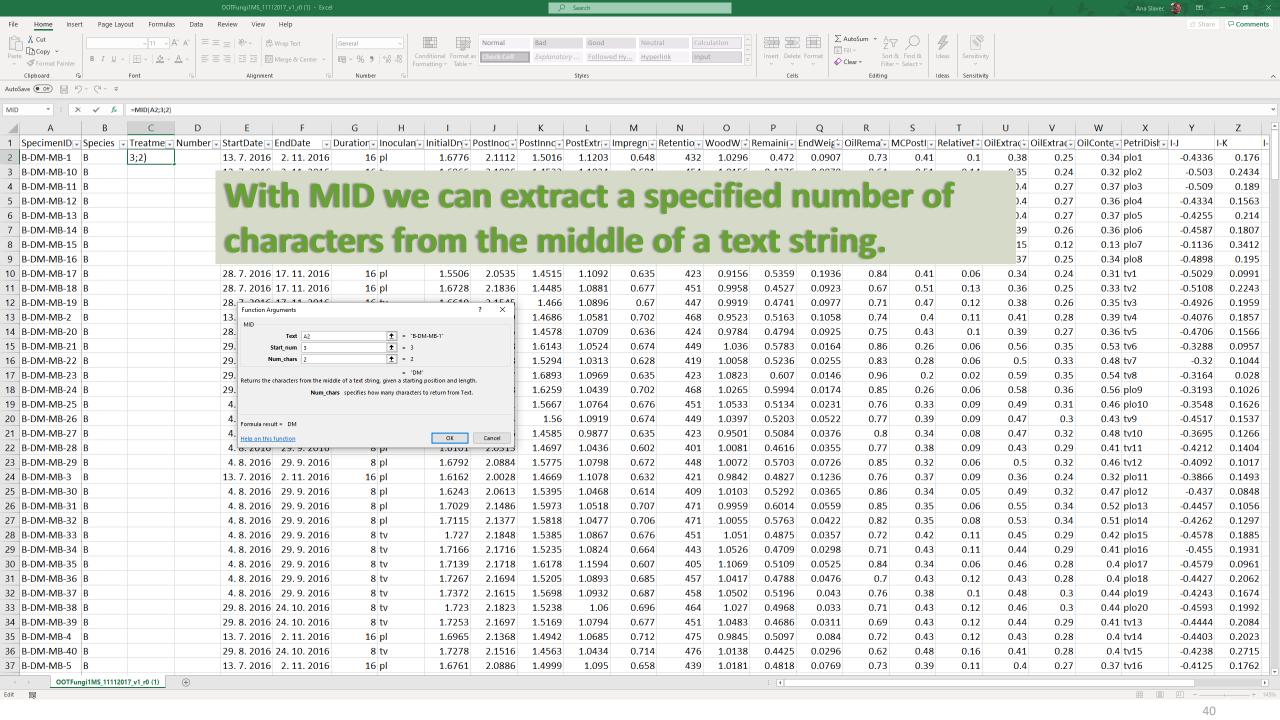


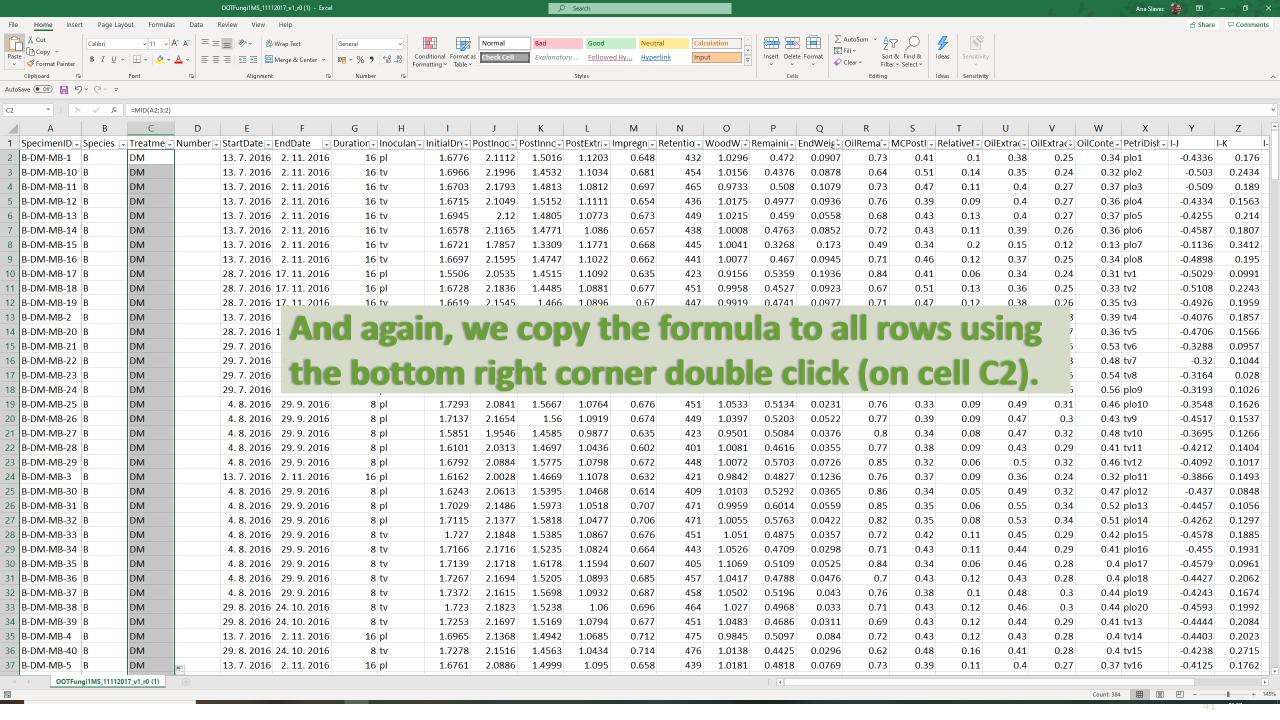


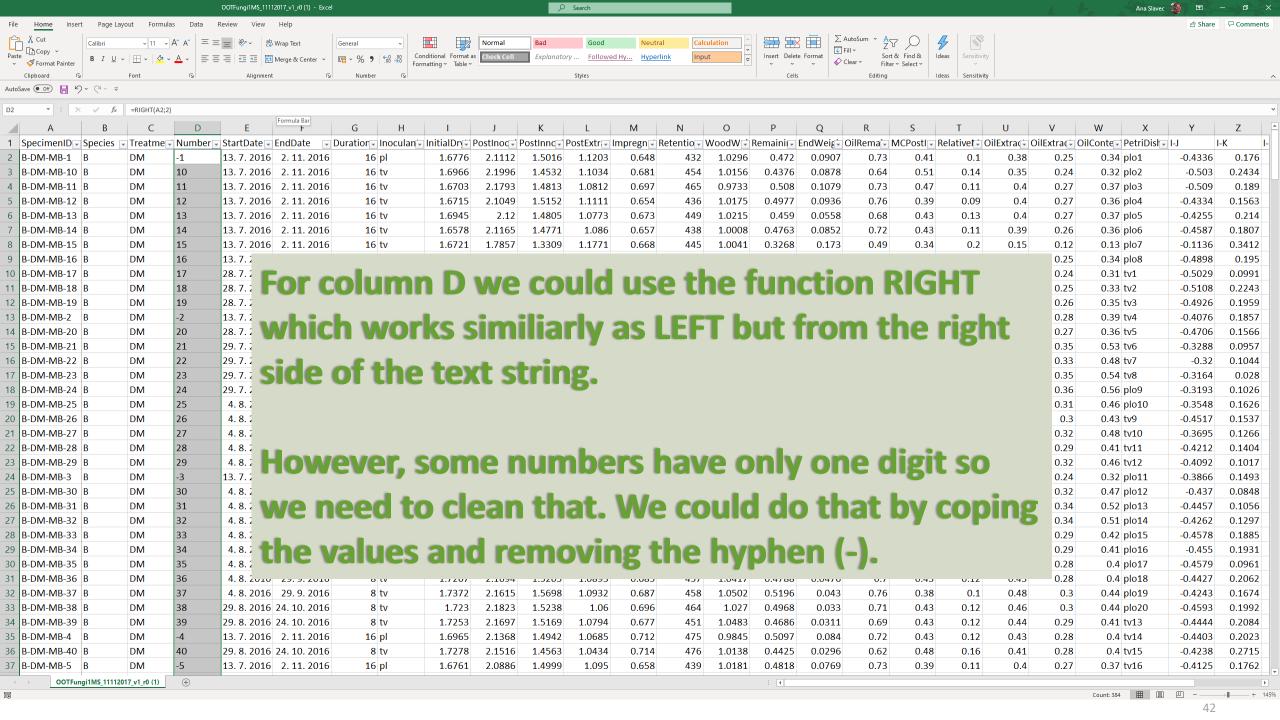


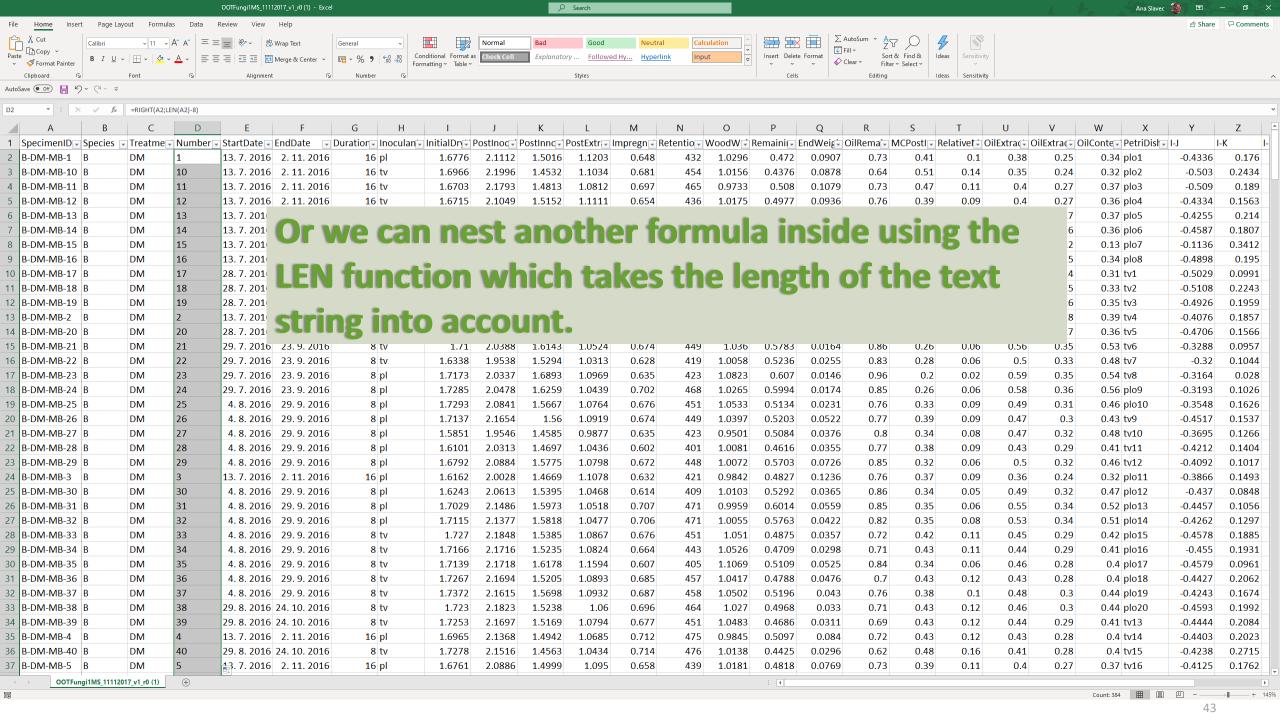


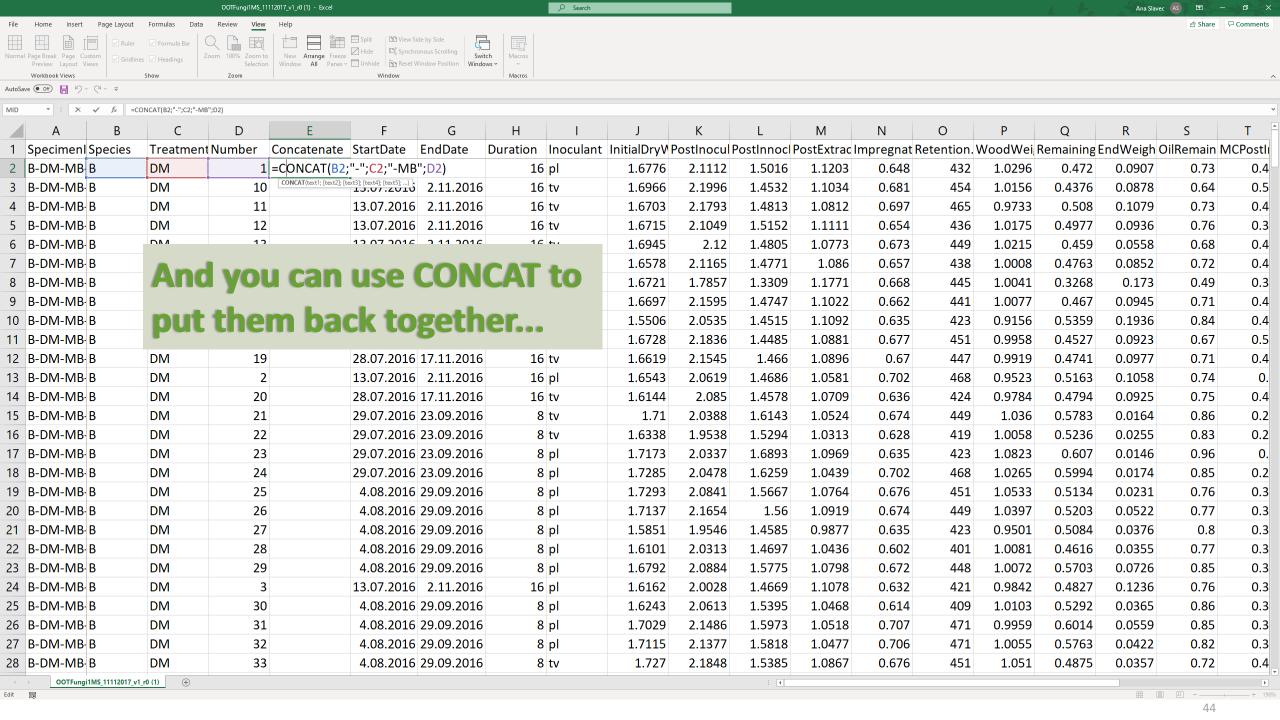


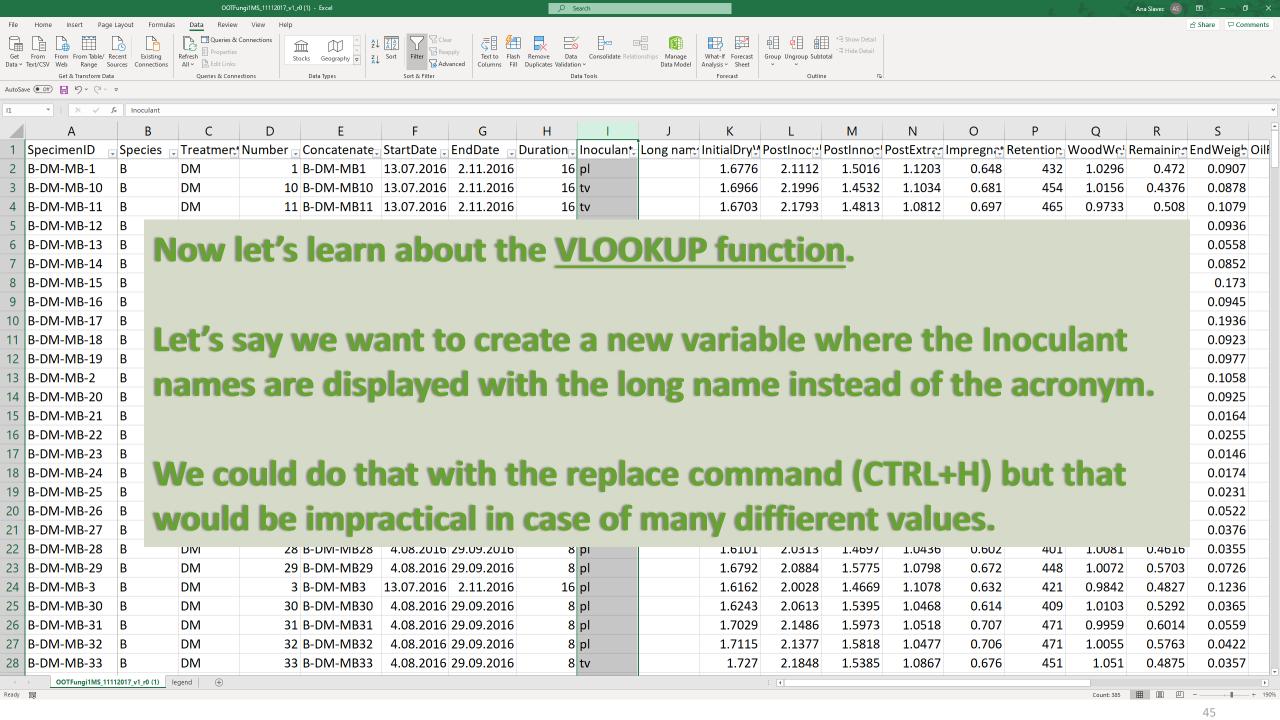


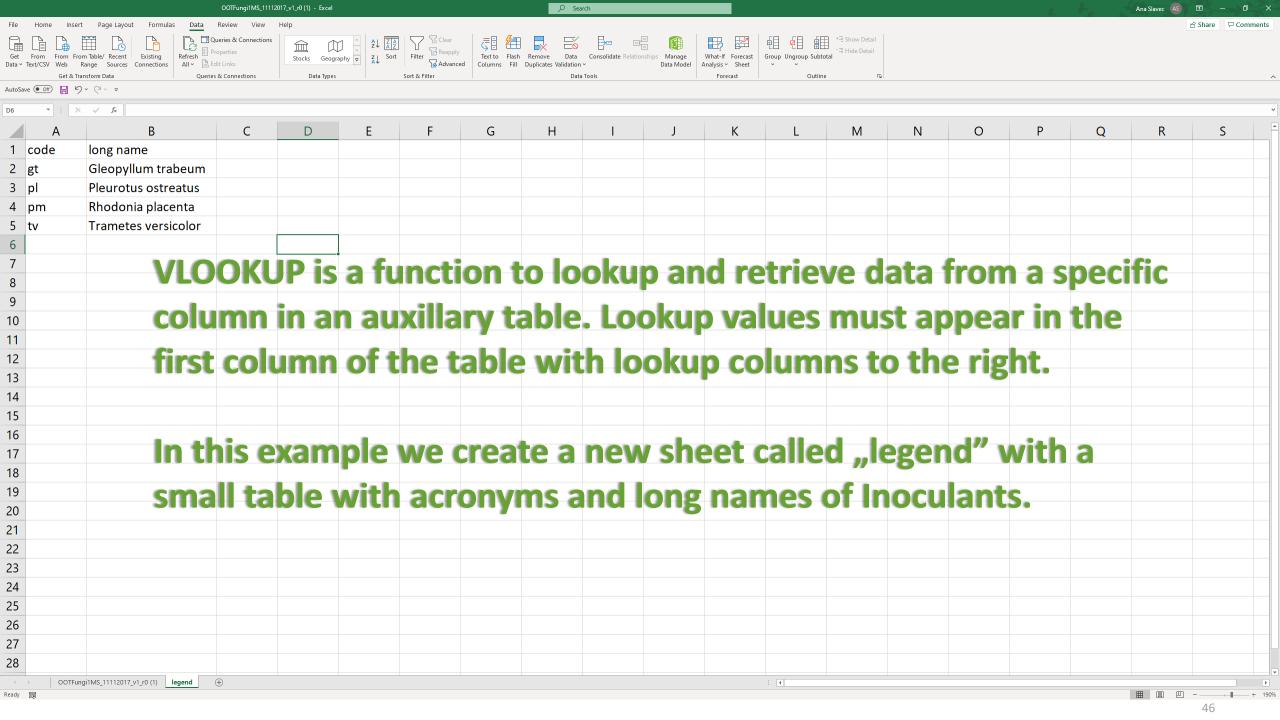


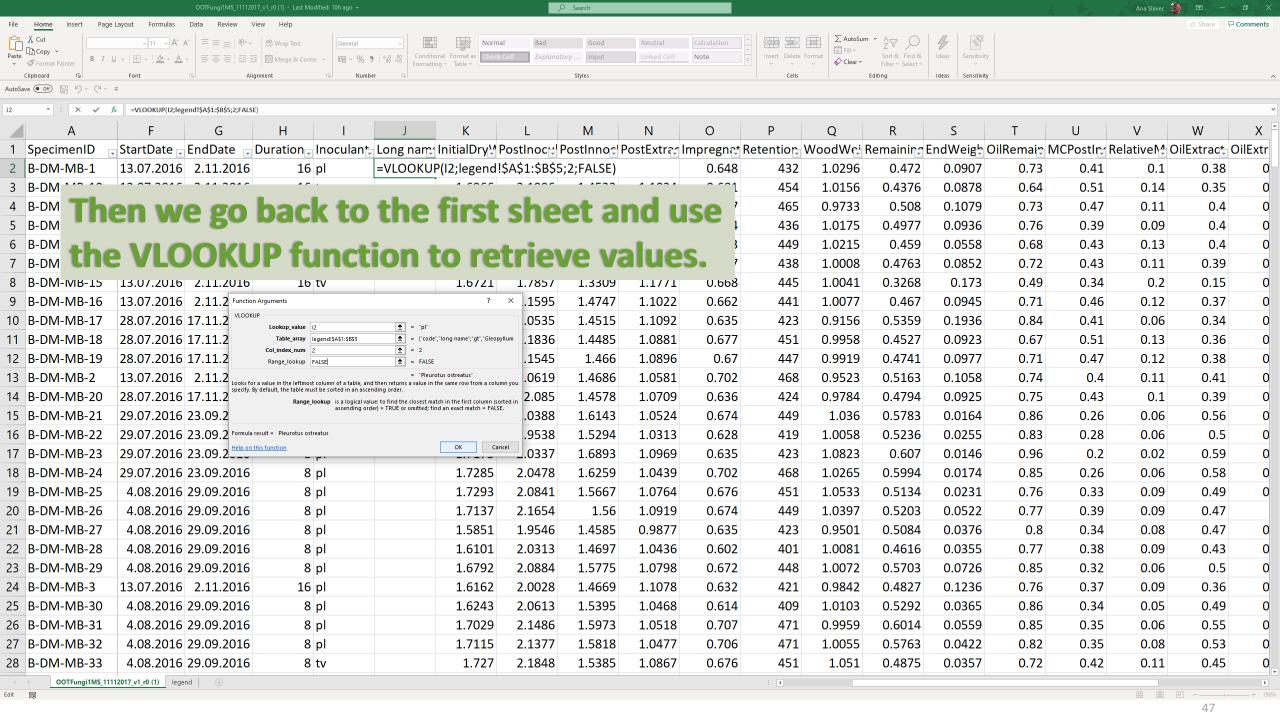


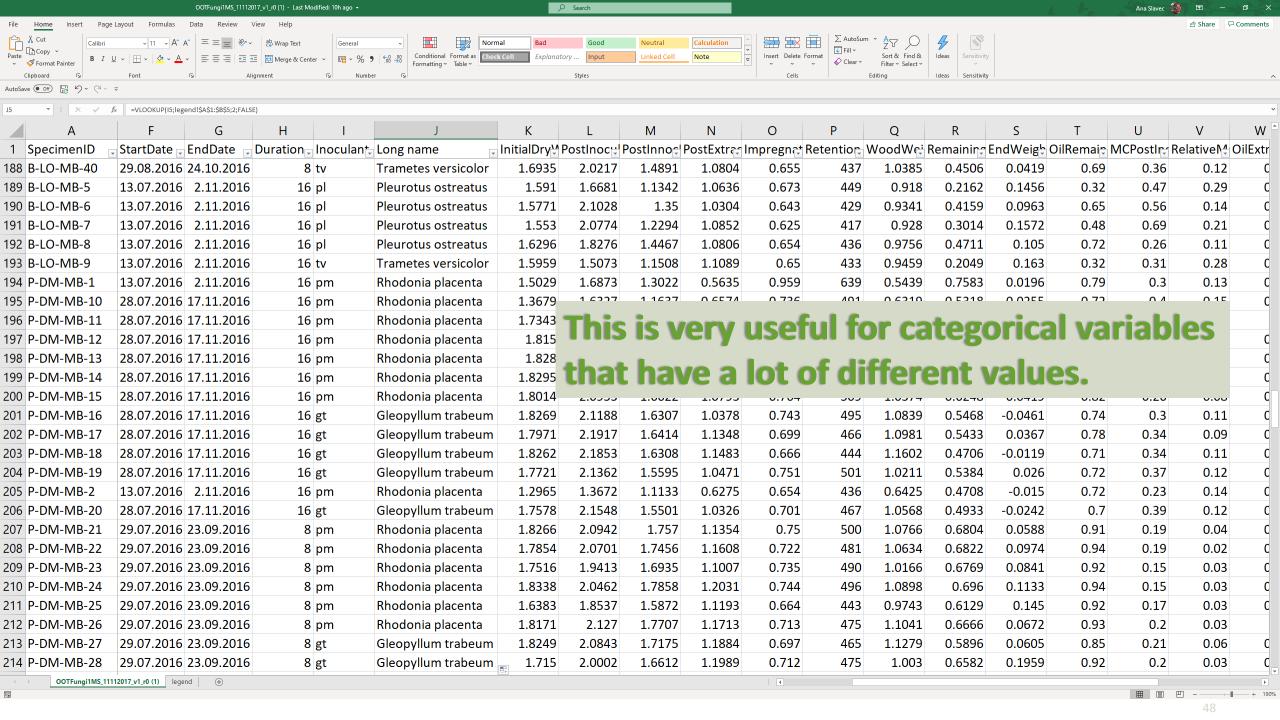


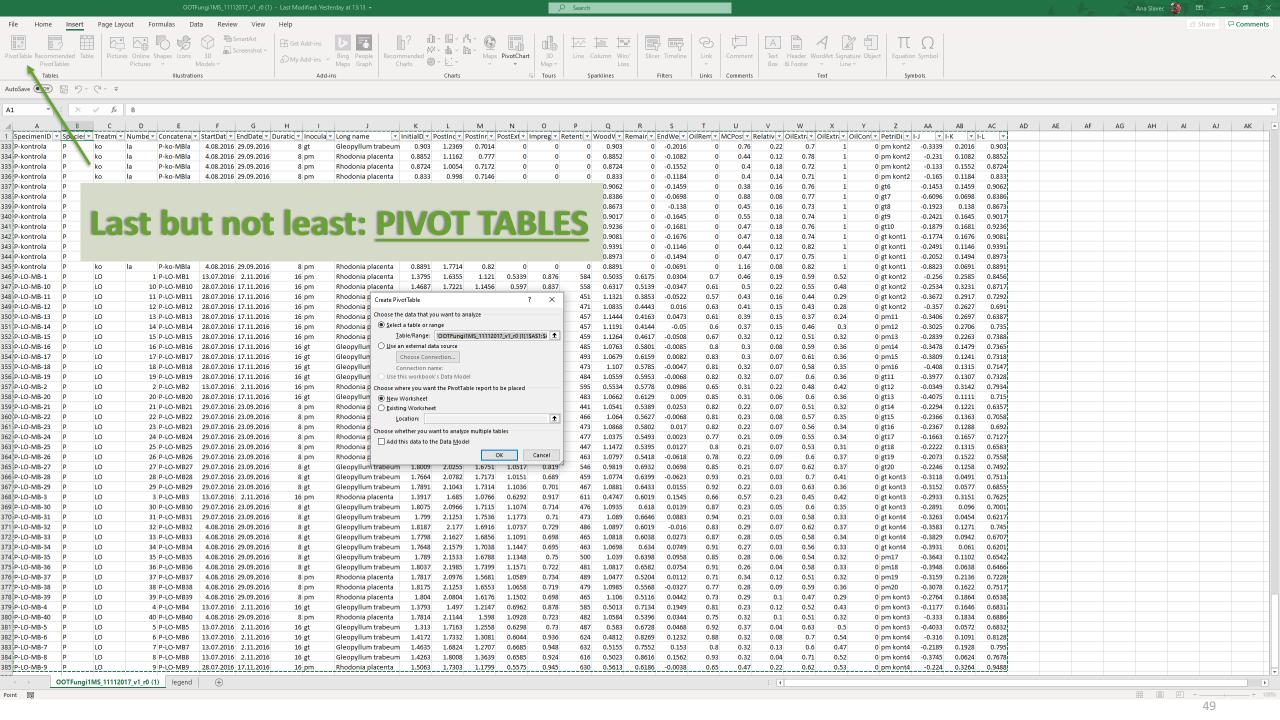


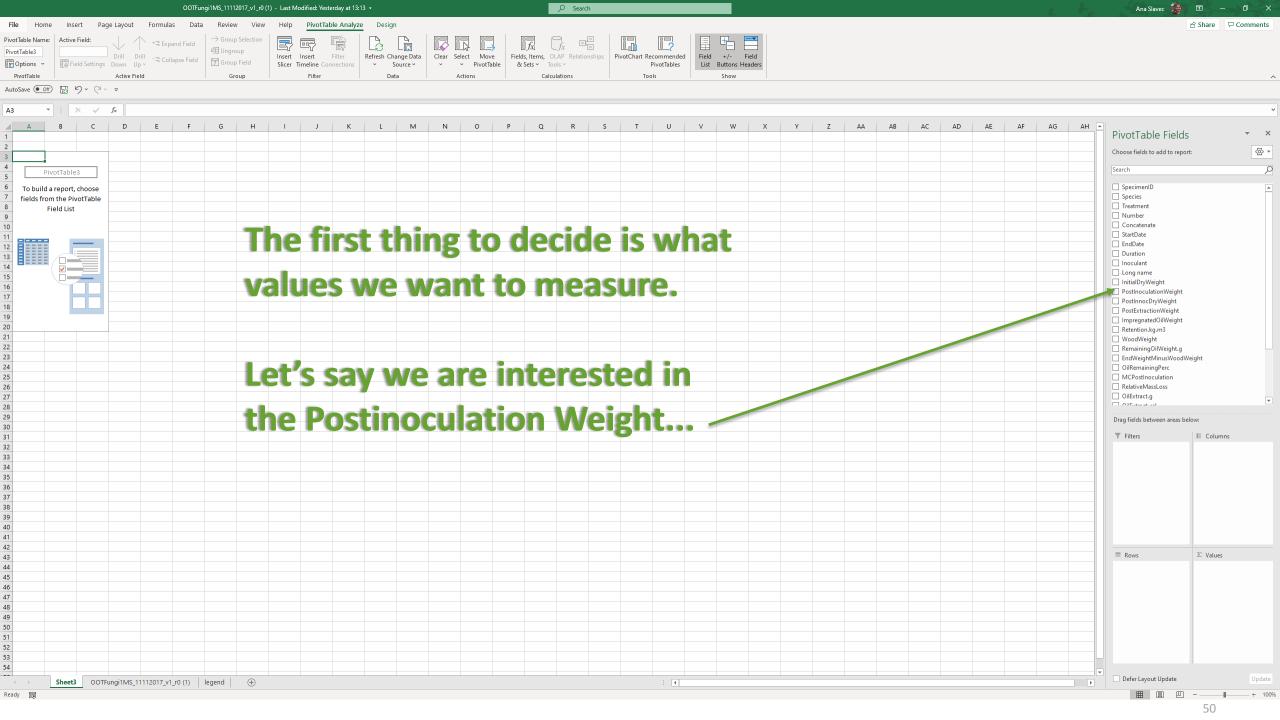


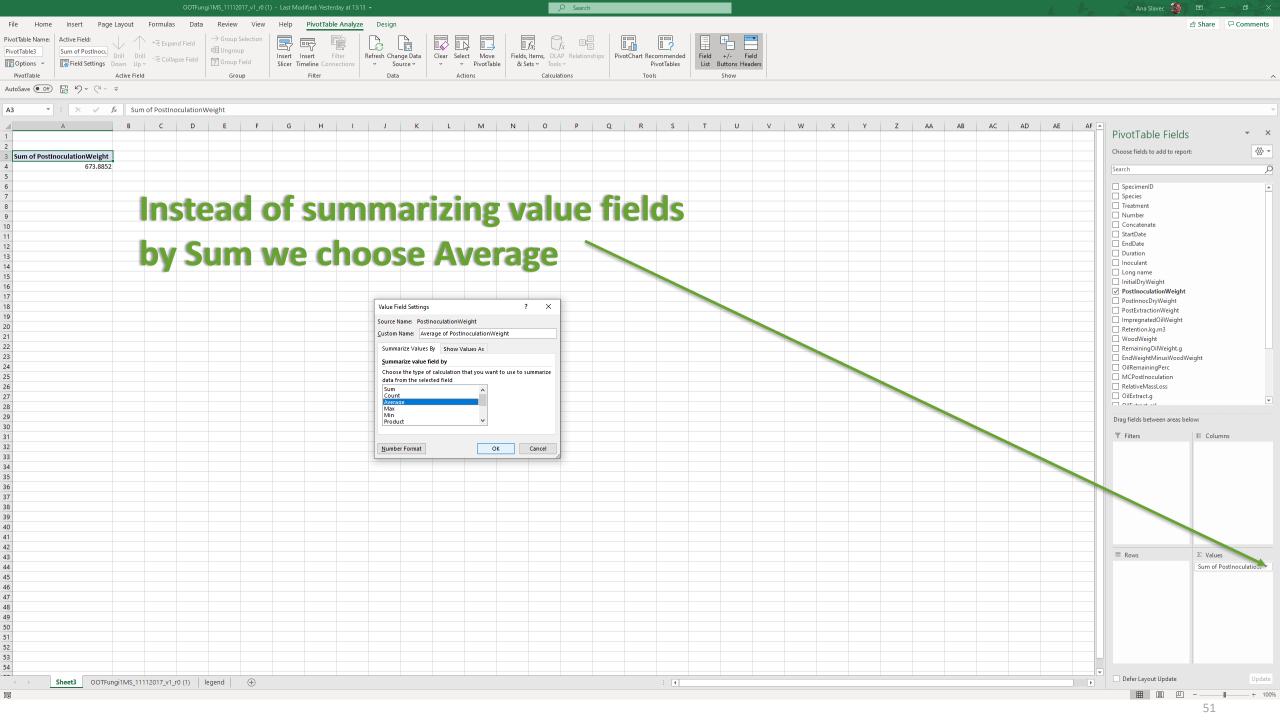


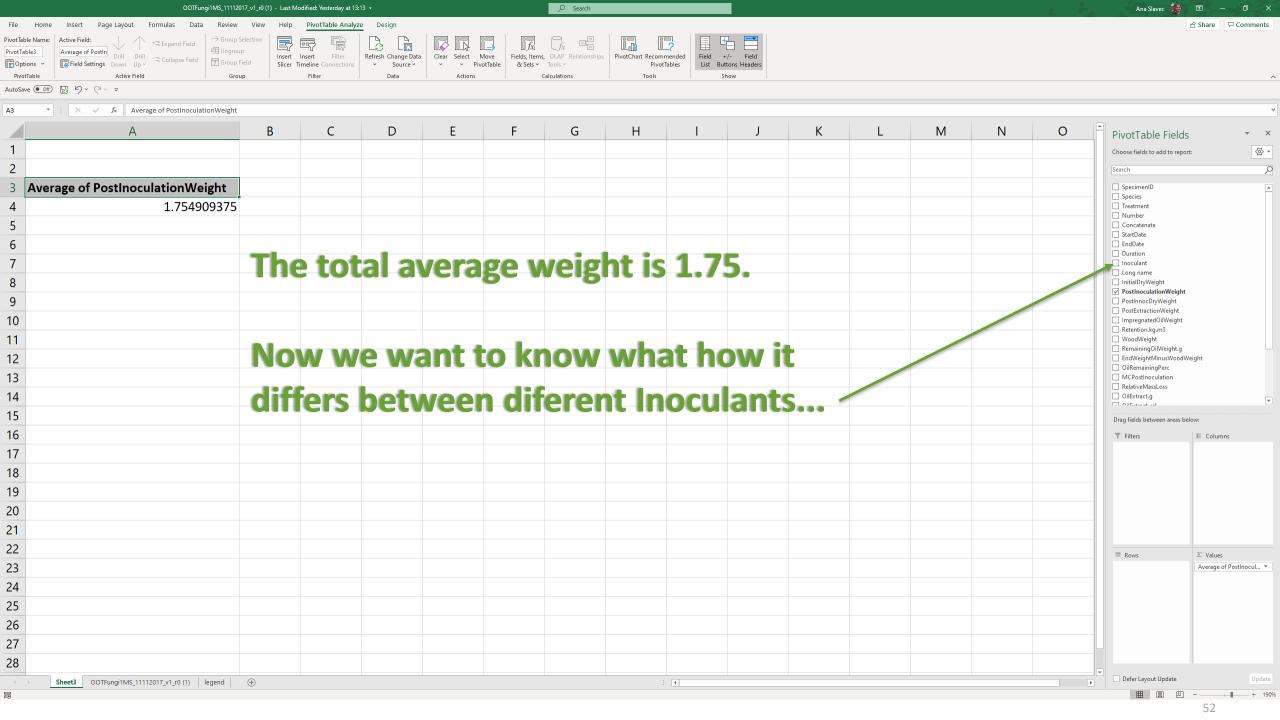


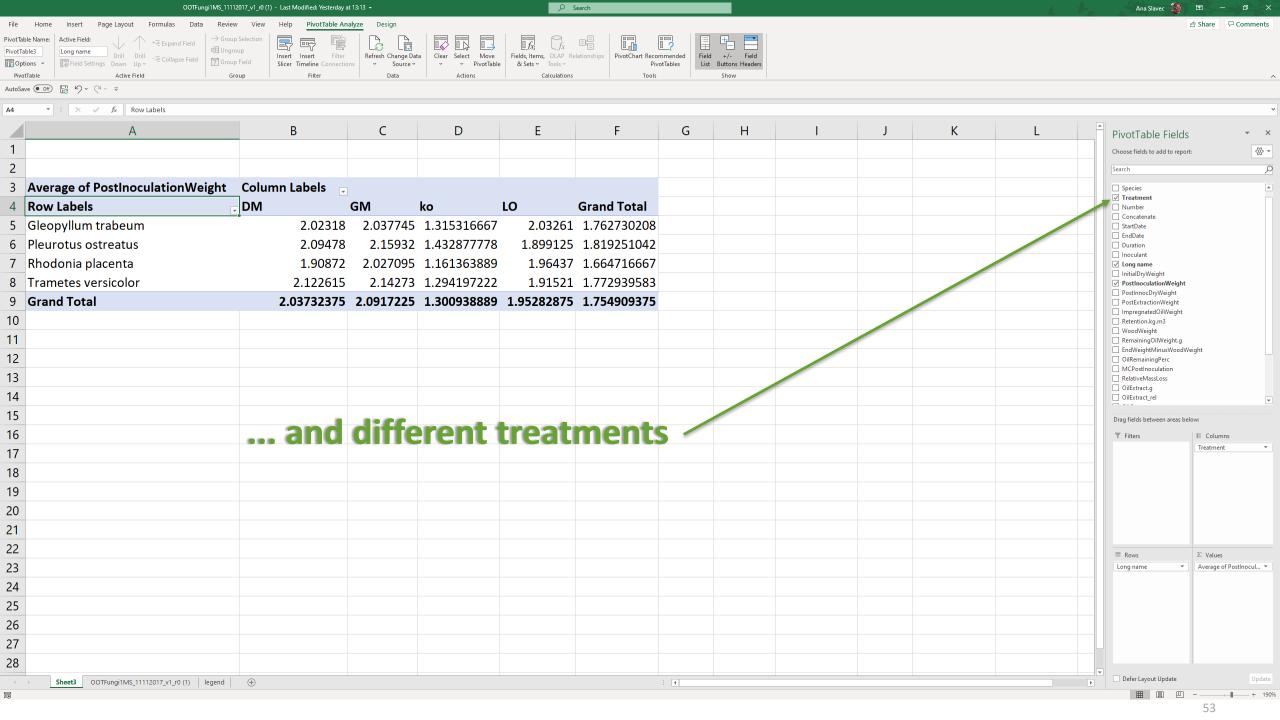














Ask me anything about Excel:)

ana.slavec@innorenew.eu