

10-8-2020

## Virtual DSS: Excelling with Excel Part 1

Molly Castro

*Florida International University*

Rebecca Bakker

*Florida International University*

Jill Krefft

*Florida International University*

Jamie Rogers

*Florida International University*

Follow this and additional works at: <https://digitalcommons.fiu.edu/dss-resources>



Part of the [Library and Information Science Commons](#)

---

### Recommended Citation

Castro, Molly; Bakker, Rebecca; Krefft, Jill; and Rogers, Jamie, "Virtual DSS: Excelling with Excel Part 1" (2020). *Digital Scholar Studio Resources*. 1.

<https://digitalcommons.fiu.edu/dss-resources/1>

This work is brought to you for free and open access by FIU Digital Commons. It has been accepted for inclusion in Digital Scholar Studio Resources by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

# EXCELLING WITH EXCEL: A TWO-PART WEBINAR SERIES

Part 1

Jill Krefft, Institutional Repository Coordinator  
[jkrefft@fiu.edu](mailto:jkrefft@fiu.edu)

Molly Castro, Digital Humanities Librarian  
[mocastro@fiu.edu](mailto:mocastro@fiu.edu)

Rebecca Bakker, Digital Collections Librarian  
[rbakker@fiu.edu](mailto:rbakker@fiu.edu)

FIU LIBRARIES' DIGITAL SCHOLAR  
STUDIO PRESENTS:

# EXCELLING WITH EXCEL

Part one of a two-part webinar series on best practices, tips, and tricks, for beginning to advanced tabular data wranglers.

**THURS OCT 8**  
**2:30-3:30**

RSVP at [go.fiu.edu/excelp1](https://go.fiu.edu/excelp1)

Data Entry

Organizing/Cleaning Data

Subsetting & Sorting

Basic Statistics

# WHAT IS EXCEL GOOD FOR?

Complicated Analysis/Statistics

Lots of data

Other tools to consider

- OpenRefine
- R
- SPSS
- SAS

**WHAT IS EXCEL  
NOT GOOD FOR?**

# EXCEL BEST PRACTICES

---

**Keep  
your  
data  
tidy**

**Put variables into one column**

---

**Put each observation in its own row**

---

**Don't combine multiple pieces of information in one cell**

---

**Be consistent with data entry**

---

# EXCEL BEST PRACTICES

---

Keep a copy of your Raw Data!

---

Create a new file or tab with your cleaned or analyzed data

---

Keep track of your analysis and changes either in a separate text file or a new tab in your workbook

---

Export cleaned data to a csv

# EXCEL DON'TS

---

**Multiple tables on one sheet**

---

**Using "0" or blanks for null values**

---

**Using Colors or other formatting to convey information**

---

**No Special characters [ \$ % , ( ) < > ~ : ]**

# LET'S TALK ABOUT DATES

13	What I typed	What my excel formats it as	DOW, Month, Day, Year	m/d/yyyy	yyyy-mm-dd	How excel stores it
14	1-Mar-19	1-Mar	Friday, March 1, 2019	3/1/2019	3/1/2019	43525
15	March 1, 1887	March 1, 1887	March 1, 1887	March 1, 1887	March 1, 1887	March 1, 1887
16						
17						

- It's safer to store dates as:
  - Month, Day and Year in separate columns (Especially if you are working with historic data before and after 12/31/1899)
  - YYYYMMDDhhmmss format (ISO 8601 – international date standard)

Choosing one of these options helps with

- Easier for sorting
- Removes ambiguity and potential for errors

The screenshot shows an Excel spreadsheet with the following columns: gene names, internal date format, and default date format. The data is as follows:

	gene names	internal date format	default date format
1	APR-1	35885	1-Apr
2	APR-2	35886	2-Apr
3	APR-3	35887	3-Apr
4	APR-4	35888	4-Apr
5	APR-5	35889	5-Apr
6	DEC-1	36129	1-Dec
7	DEC-2	36130	2-Dec
8	DEC1	36129	1-Dec
9	DEC2	36130	2-Dec
10	MAR1	35854	1-Mar
11	MAR2	35855	2-Mar
12	MAR3	35856	3-Mar
13	NOV1	36099	1-Nov
14	NOV2	36100	2-Nov
15			

Zeeberg, B.R., Riss, J., Kane, D.W. *et al.* Mistaken Identifiers: Gene name errors can be introduced inadvertently when using Excel in bioinformatics. *BMC Bioinformatics* 5, 80 (2004). <https://doi.org/10.1186/1471-2105-5-80>



# DATES

Again, dates in Excel can be frustrating!

To customize your dates:

Open the dialog box Custom Number (right click-->format cell)

In this dialog box, you select '**Custom**' in the Category list and write the date format code in '**Type.**' Here we are using:

**'yyyy-mm-dd'**

Format Cells

Number Alignment Font Border Fill Protection

Category: Sample

2009-01-05

Type: yyyy-mm-dd

\_\$\* #,##0.00\_);\_(\$\* (#,##0.00);\_(\$\* "-"??\_);\_(@\_)  
\_(\* #,##0.00\_);\_(\* (#,##0.00);\_(\* "-"??\_);\_(@\_)  
"Yes";"Yes";"No"  
"True";"True";"False"  
"On";"On";"Off"  
[\$€-x-euro2] #,##0.00\_);[Red]([\$€-x-euro2] #,##0.00)  
[\$-en-US]dddd, mmmm d, yyyy  
[\$-en-US]h:mm:ss AM/PM  
yyyy-dd-mm  
mmm-yyyy  
yyyy-mm-dd

Delete

Type the number format code, using one of the existing codes as a starting point.

Cancel OK

# DATES



	A	B	C
1	Date	Date	Date
2	17-Nov-72	1972-11-17	1972-11-17
3	4-Dec-72	1972-12-04	1972-12-04

	A	B	C
1	Date	Date	Date
2	17-Nov-72	1972-11-17	1972-11-17
3	4-Dec-72	1972-12-04	1972-12-04

Number Alignment Font Border Fill Protection

Category:

- General
- Number
- Currency
- Accounting
- Date
- Time
- Percentage
- Fraction
- Scientific
- Text**
- Special
- Custom

Sample

Date

Text format cells are treated as text even when a number is in the cell. The cell is displayed exactly as entered.

# SPACES

While spaces seem harmless, and maybe even INVISIBLE, they may cause errors in your formulas. We present two easy ways to remove extra spaces: leading, trailing, and in-between.

	A
1	cat
2	cat
3	cat in the hat
4	

**TRIM function:** Formula that removes all spaces from a text string except for single spaces between words. **=TRIM(cell value)**

**Find and Replace:** A simple way to get rid of spaces quickly, especially if the problem repeats throughout your data with the same “type” of space.

# REMOVING TEXT STRINGS (OR FILE EXTENSIONS)

Sometimes, you may need to trim specific text from an entry in excel. In our example, we will use file extensions.

This can be done with the LEFT function. If you have a long filename and don't feel like counting, use it with FIND:

**=LEFT(cell value,FIND(".",cell value)-1)**

To make the formula more versatile:

- "." in the formula can be replaced with any delimiter (comma, underscore, etc.)
- LEFT can be replaced with RIGHT

LEFT Function: extracts a given number of characters from the left side of a supplied text string. For example, LEFT("document.txt",8) returns "document".

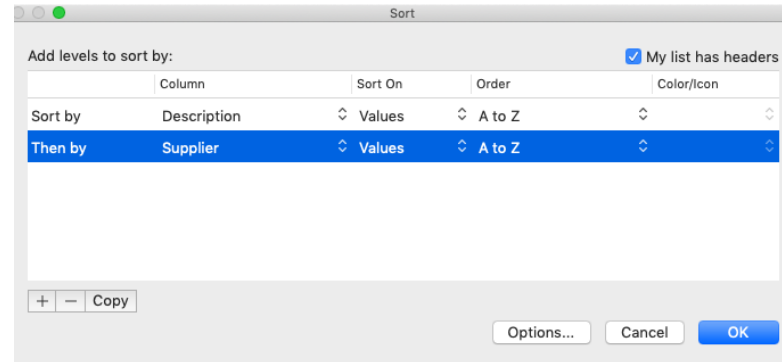
FIND: returns the position (as a number) of one text string inside another.


# SORT & FILTER

Sort and filter are great ways to do quick analysis of your data, see patterns, and pull out specific information.

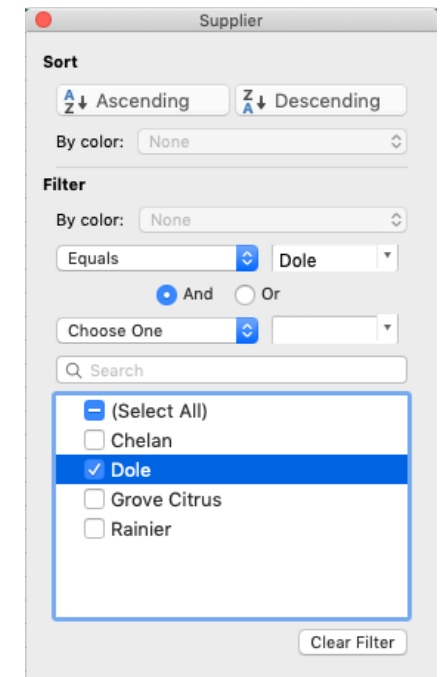
To get the most out of these functions:

- Have clean data
- Avoid blank columns or rows



**Sorting** can be done by 'quick sort'  buttons or by opening the sort dialogue. This enables you to add levels by which to sort more granularly.

When turned on, **filter** arrows appear on each column header. When clicked, a list of unique values appears with tick boxes that can be utilized to show or hide certain data.



# LOWERCASE/UPPERCASE

To clean up text, names, titles, etc. it helps to have consistent cases.

This can be done with the following functions:

**=UPPER(cell value)**

**=LOWER(cell value)**

**=PROPER(cell value)**

UPPER

lower

Proper

**Alternatively, you can make the whole cell lowercase, and have the first letter replaced with a capitalized letter:**

**=REPLACE(LOWER(cell value),1,1,UPPER(LEFT(cell value,1)))**

# SPECIAL CHARACTERŚ

Text may need to be cleaned of characters that are either there accidentally or are problematic for certain computer systems that are reading your data.

This can be helpful for:

- Diacritics
- Symbols
- Spaces

```
=SUBSTITUTE(cellvalue,"thing  
you want replaced","thing  
you want to replace it with")
```

It's like Find and Replace but  
targeted to specific cells!

# CONCATENATING CELLS

When your text is in separate cells and you want to bring them together – concatenate!

Use:

`=CONCATENATE(cell value,cell value)`

Or:

`=(cell value)&(cell value)`



**"Add additional text by using quotes!"**



# TEXT TO COLUMNS

When your text is in the same cell and you want to split it up or otherwise reorganize it – text to columns!

Use the Text to Columns Wizard to walk through each step.

Convert Text to Columns Wizard - Step 1 of 3

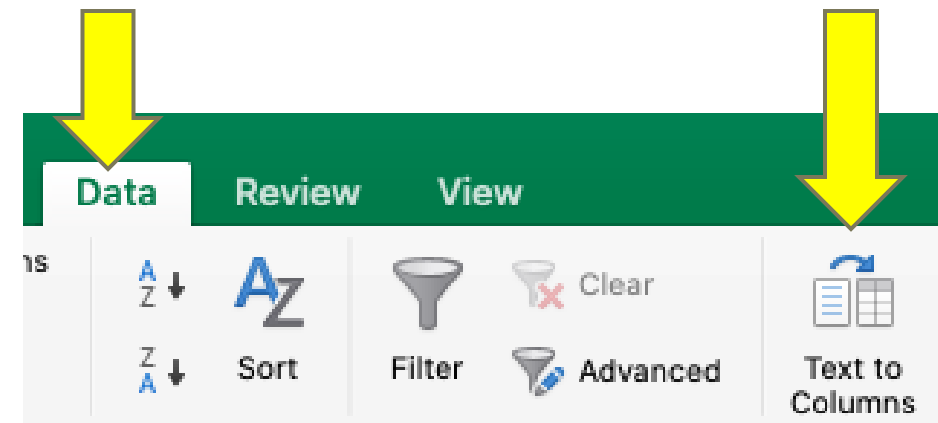
The Text Wizard has determined that your data is Fixed Width.  
If this is correct, choose Next, or choose the Data Type that best describes your data.

Delimited - Characters such as commas or tabs separate each field.  
 Fixed width - Fields are aligned in columns with spaces between each field.

Preview of selected data:

Preview of selected data:	
7	Automotive repair shops - Automobiles - Central business districts - Cassandra Hotel
8	Ships - Shipwrecks - Coquimbo (Barque)
9	Anniversaries - Birthday cakes - Bakers - Lucy's Bakery
10	Central business districts - Automobiles - Traffic signs and signals - Stores - Retail
11	Marina - Boats and boating - Restaurants - Two Georges Waterfront Café
12	Woodbright Road
13	Motion picture theaters - Automobiles - Stepped gables
14	Anniversaries - Birthday cakes - Children
15	Anniversaries - Birthday cakes - City council members

Cancel < Back **Next >** Finish



# JOIN US FOR PART 2!

Part 2 will build on part one and include more advanced formulas and concepts, such as working with hyperlinks, conditional formulas, and using Excel with external tools.

Anything else you would like us to cover?  
Let us know!

Register at [go.fiu.edu/excelp2](https://go.fiu.edu/excelp2)

FIU LIBRARIES' DIGITAL SCHOLAR  
STUDIO PRESENTS:

# EXCELLING WITH EXCEL

Part two of a two-part webinar series on best practices, tips, and tricks, for beginning to advanced tabular data wranglers.

**THURS NOV 5**  
**2:30-3:30**

RSVP at [go.fiu.edu/excelp2](https://go.fiu.edu/excelp2)

# COMING UP NEXT IN THE DSS

This introductory webinar will cover basic concepts, vocabulary, methods, and tools for getting started with text analysis in the humanities. No coding experience is required. Participation is open to scholars from any discipline, though concepts will be presented with a humanities focus.

**Wednesday, Oct 21st 2:30-3:30pm**

**Register at [go.fiu.edu/textanalysis](https://go.fiu.edu/textanalysis)**

FIU Libraries' Digital Scholar Studio Presents

## INTRODUCTION TO TEXT ANALYSIS

**for the Humanities**

**Wednesday, 10/21  
2:30-3:30pm**

Register at  
[go.fiu.edu/textanalysis](https://go.fiu.edu/textanalysis)

This introductory webinar will cover basic concepts, vocabulary, methods, and tools for getting started with text analysis in the humanities. **No coding experience is required.** Participation is open to scholars from any discipline, though concepts will be presented with a humanities focus.