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.
EXCELING WITH EXCEL: A TWO-PART WEBINAR SERIES

Part 1

Jill Krefft, Institutional Repository Coordinator
jkrefft@fiu.edu

Molly Castro, Digital Humanities Librarian
mocastro@fiu.edu

Rebecca Bakker, Digital Collections Librarian
rbakker@fiu.edu

FIU LIBRARIES' DIGITAL SCHOLAR STUDIO PRESENTS:

EXCELING 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
WHAT IS EXCEL GOOD FOR?

- Data Entry
- Organizing/Cleaning Data
- Subsetting & Sorting
- Basic Statistics

Adapted from The Carpentries. Licensed under CC-BY 4.0 2018–2020
WHAT IS EXCEL NOT GOOD FOR?

Complicated Analysis/Statistics

Lots of data

Other tools to consider
- OpenRefine
- R
- SPSS
- SAS
### Excel Best Practices

<table>
<thead>
<tr>
<th>Keep your data tidy</th>
<th>Put variables into one column</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Put each observation in its own row</td>
</tr>
<tr>
<td></td>
<td>Don't combine multiple pieces of information in one cell</td>
</tr>
<tr>
<td></td>
<td>Be consistent with data entry</td>
</tr>
</tbody>
</table>

Adapted from The Carpentries Licensed under CC-BY 4.0 2018–2020
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

Adapted from The Carpentries Licensed under CC-BY 4.0 2018–2020
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

• 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


Adapted from The Carpentries Licensed under CC-BY 4.0 2018–2020
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'
### Dates

#### Me:
- 12
- 5
- 12.5
- 1/12/1900
- 12:00:00 PM

#### Excel:
- 12
- 5
- 12

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Date</td>
<td>Date</td>
</tr>
<tr>
<td>2</td>
<td>17-Nov-72</td>
<td>1972-11-17</td>
</tr>
<tr>
<td>3</td>
<td>4-Dec-72</td>
<td>1972-12-04</td>
</tr>
</tbody>
</table>

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

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.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cat</td>
</tr>
<tr>
<td>2</td>
<td>cat</td>
</tr>
<tr>
<td>3</td>
<td>cat in the hat</td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

**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.
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 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

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

This can be done with the following functions:

\[ =\text{UPPER}(\text{cell value}) \]
\[ =\text{LOWER}(\text{cell value}) \]
\[ =\text{PROPER}(\text{cell value}) \]

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

\[ =\text{REPLACE}(\text{LOWER}(\text{cell value}), 1, 1, \text{UPPER}(\text{LEFT}(\text{cell value}, 1))) \]
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.
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
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