Mopping Up Metadata
Excel Functions, Formatting, Formulae

AABC-ARMA VI “Dominion of the Digital” Conference
Victoria, BC, 28 April 2017

Katrina Cohen-Palacios, Adjunct Archivist
Clara Thomas Archives and Special Collections
York University Libraries
map

» formatting, conditional

» formulae: concatenate; len; countif; today, date, year

» functions: filters, sort, data validation
conditional formatting

Difficult to spot errors in data input without conditional formatting

Conditional formatting visualizes, highlights and categorizes error, including invisible errors such as double spaces (in blue), duplicate data (in purple), and character limits (in orange).
conditional formatting

» visualize data
» highlight information
» customize with formulae

potential applications

» mark required fields in red
» highlight common errors (restricted characters, double spaces, space at end), duplicates, etc.
conditional formatting

» double spaces

» leading spaces

» trailing spaces

» character limits
conditional formatting

rules manager

» set rule precedence

Learn more at: https://support.office.com/en-us/article/Manage-conditional-formatting-rule-precedence-06aeeb54-0fad-4b2d-a43b-6db6f6f3ca26
conditional formatting

pitfalls

» conditional formatting removed when pasting information to cells

» solutions:

1. special paste values into cell & leave first row blank to special paste formatting to completed data

2. use rules manager to visually verify the application of formatting to a set of rows and/or columns (see previous slide)
formulae

» concatenate
» len
» countif
» today, date, year
formulae » concatenate

» strings two or more text into one string

=CONCATENATE (A1, “text”, B1, …)

» use quotation marks to insert text

» concatenate replaced by concat in Excel 2016

Learn more at: https://support.office.com/en-us/article/CONCATENATE-function-8f8ae884-2ca8-4f7a-b093-75d702bea31d
formulae » concatenate » uses

potential applications

» determine characters used with LENS formula
» automate fields and standardized notes

pitfalls

» if referenced cells are deleted, the information is also deleted from formulated cell
» prior to deleting referenced cell, copy and special paste values in formulated cell
formulae » **concatenate** » spill cells

» opposite: “text to columns” option on “data” tab
formulae » len

» returns the number of characters in a text string

=LEN(A1+B1)

pitfall

» number and date formatting is not included in count
   (ie: 100 formatted as $100.00 returns 3)

potential applications

» calculate characters based on restrictions
LEN formula combined with CONCATENATE to determine character limits for a file path in Sharepoint
formulae » **countif**

» calculates cells that meet a criterion

`=COUNTIF(A2:A99, "text")`

`=COUNTIF(A2:A99, A1)`

» can use wildcard characters (?, *) in criteria

» remove leading spaces, trailing spaces, etc

» **COUNTA** counts cells that are not empty

» **COUNTBLANK** counts empty cells
formulae » today, year, day

» =TODAY() returns the current date
» =TODAY()+5 returns date plus five days
» =YEAR(TODAY()) returns year of current date
» =MONTH(TODAY()) returns month of date
» =DAY(TODAY()) returns day of current month

potential applications: calculate anniversaries

=YEAR((TODAY())-YEAR(C1))

pitfall: cannot calculate dates prior to 1900
functions

» filters

» sort

» data validation
filter
**filter**

hide unwanted data (but does not delete it)

» auto-filters: by value, by format, or by criteria

» each filter type is mutually exclusive

» customize filters

**common error**: ensure all columns and rows are included in the filtered data

**tip**: use freeze pane to view header columns when scrolling; boolean searching
filter » uses

potential applications

» verify consistency/format of data
» highlights misspellings
» used to inventory vendor outputs
  ex: # of rows = # of PDFs returned
» view upcoming anniversary dates

pitfall: filtered out data is deleted when file is saved as a CSV
Quickly spot inconsistent data formatting

Discover and calculate data such as upcoming anniversaries for social media posts
sort
sort

rearrange data

» by text: A to Z or reverse

» by numbers: smallest to largest or reverse

» by date/time: oldest to newest or reverse

common errors: remove leading spaces, verify if all cells are formatted text or numbers, sorting does not move hidden rows
sort » uses

potential applications

» rearrange data to reflect original order by sorting by box, file, and item numbers

» review entire series and maintain original order

» spot inconsistencies in data input
data validation

restrict type of data/values entered in a cell

» set up a control list in a worksheet and select the entire column to ensure additions/deletions are updated automatically the validation list

Image from: https://support.office.com/en-us/article/More-on-data-validation-f38dee73-9900-4ca6-9301-8a5f6e1f0c4c
data validation » settings
data validation » settings

error alert appears if user enters invalid data
  » stop: user must retry or cancel data input
  » warning: warning appears but does not prevent
  » information: informs users data is invalid
data validation » settings

input message: define instructions
data validation » pitfalls

» pasting data removes data validation

» paste values only via “special paste” option

» error alerts are not applied retroactively

» circle invalid data to review
Data validation uses a control list from the spreadsheet file.

Data validation creates drop-down list for input.

Data validation circles invalid data and creates input warnings.
data validation » pitfalls

multiple users

» create control list as a separate workbook
» design name for each list from formulas tab
» in data validation workbook, add the list via the name manager and enter address in “refers to” section of data validation
questions?

conditional formatting
  » visualize data

filter » hide data

sort » rearrange data

data validation
  » control data entered in cell

formulae
  =CONCATENATE (A1, “text”)
  =LEN(A1+B1)
  =COUNTIF(A2:A99,”TEXT”)
  =YEAR((TODAY())-YEAR(C1))

contact info
  kcohenp@yorku.ca
  twitter: @katrina_cp