

Keep @ Downsvie Metadata Matching Instructions

August 20, 2019

Start with clean data files for WL and UTL.

- OCLC and ISBN numbers must be on individual rows, extraneous data removed
- Insert a letter a in front of ISBN and OCLC numbers on all the sheets (ASAP Utilities)
- clear leading, trailing and excessive spaces from ISBN and OCLC columns
- clear all formats from ISBN and OCLC columns

Prepare a spreadsheet with the following information

Filename/filepath: U:\...\2-Metadata Matching in Progress\...\WL-UTLdata.xlsx

Sheet_name: WL-ISBN (Clean WL data) Headings: CALLNUM-WL, BIBNUM, BARCODE, ITEMNUM, TITLE-WL, ISBN
Sheet_name: WL-OCLC (clean WL data) Headings: CALLNUM-WL, BIBNUM, BARCODE, ITEMNUM, TITLE-WL, OCLC,
Sheet_name: UTL-ISBN (clean UTL data) Headings: UTLLIB, CATKEY, CALLNUM-UTL, CALLNUM2-UTL, TITLE-UTL, ISBN <ul style="list-style-type: none">• remove any rows that do not contain ISBNs
Sheet_name: UTL-OCLC (clean UTL data) Headings: UTLLIB, CATKEY, CALLNUM-UTL, CALLNUM2-UTL, TITLE-UTL, OCLC <ul style="list-style-type: none">• remove any rows that do not contain OCLC numbers

Run the Python script

Save DVmatch.py to your computer.

Open IDLE. File > open > DVmatch.py

Within the script, change file paths (4 of them, all colour-coded green) to locate your file. Also add the file path to where you would like to save the output document. (WL-UTL matches.xlsx). Save these changes

Example: 'U:\...\WL-UTL data.xlsx'

On this file: run > run module

Open and clean up WL-UTL data.xlsx file

Sheet_name: OCLC (This sheet contains all matches based on OCLC numbers)
Sheet_name: ISBN (This sheet contains all matches based on ISBN numbers)
Sheet_name: COMBINED (This sheet contains matches on OCLC and ISBN data combined) <ul style="list-style-type: none">• Remove all duplicates [Match on: BIBNUM, BARCODE, ITEMNUM, TITLE-WL, UTLLIB, CATKEY, TITLE-UTL] <ul style="list-style-type: none">• Conditional formatting to highlight duplicates in the ITEMNUM column, then custom sort ITEMNUM by colour on top & ITEMNUM by A-Z• Manually remove remaining duplicates, favouring downsvie locations• Doublecheck that TITLE-WL and TITLE-UTL match. If not, remove row• Move item numbers to individual rows
Sheet_name: UNMATCHED

This sheet contains all items from the original WL CLEAN file.

- Matched items need to be removed manually:
 - Filter the results to delete all rows with CATKEY information
- Remove duplicates

[Match on: TITLE-WL, BIBNUM, CALLNUM-WL, BARCODE, ITEMNUM]

- Move item numbers to individual rows

Quality Check

Ensure ITEMNUM from COMBINED and UNMATCHED align with original pull list.

- Remove any ITEMNUM not in original pull list
- Compare ITEMNUM in COMBINED and UNMATCHED to the original pull list
- Volumning from item records can be added to this data (Export from Sierra)

Create file to use as pull list, sorted into streams

Filename: [original file name](K@D by stream).xlsx

Sheet_name: DISCARD STREAM

- Filter WL-UTL Matches, COMBINED sheet on UTLLIB 'contains Downsview'

Sheet_name: STREAM 1 - UTL RECORD

- Filter WL-UTL Matches, COMBINED sheet on UTLLIB 'does not contain Downsview'

Sheet_name: STREAM 2 - NO UTL

- WL-UTL Matches, UNMATCHED sheet