

Keep @ Downsview 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
- clear leading, trailing and excessive spaces from ISBN and OCLC columns
- clear formats from ISBN and OCLC columns

Prepare a spreadsheet with the following information

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

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

Run the Python script

Download DVmatch.py

Open IDLE. file > open > DVmatch.py

Within the script, change file paths to locate your file. Save

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

On this file: run > run module

Open and clean up WL-UTL data.xlsx file

Sheetname – OCLC (This sheet contains all matches based on OCLC numbers)
Sheetname - ISBN (This sheet contains all matches based on ISBN numbers)
Sheetname – 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 downsview locations• Doublecheck that TITLE-WL and TITLE-UTL match. If not, remove row• Move item numbers to individual rows
Sheetname - UNMATCHED This sheet contains all items from the original WL CLEAN file. <ul style="list-style-type: none">• Matched items need to be removed manually:<ul style="list-style-type: none">◦ Filter the results to delete all rows with CATKEY information

- | |
|--|
| <ul style="list-style-type: none">• Remove duplicates <p>[Match on: TITLE-WL, BIBNUM, CALLNUM-WL, BARCODE, ITEMNUM]</p> <ul style="list-style-type: none">• 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

Sheetname: DISCARD STREAM
<ul style="list-style-type: none">• Filter WL-UTL Matches, COMBINED sheet on UTLLIB 'contains Downsvie'
Sheetname: STREAM 1 - UTL RECORD
<ul style="list-style-type: none">• Filter WL-UTL Matches, COMBINED sheet on UTLLIB 'does not contain Downsvie'
Sheetname: STREAM 2 - NO UTL
<ul style="list-style-type: none">• WL-UTL Matches, UNMATCHED sheet