

# Using Alma Analytics to Facilitate the Library Weeding Project

Li Chen (lchen12@kennesw.edu)  
Jason Penwell (jpenwell@Kennesaw.edu)  
Kennesaw State University Library System  
May, 2019



**KENNESAW STATE**  
UNIVERSITY  
LIBRARY SYSTEM

# Introduction

- The consolidation of Horace W. Sturgis Library and Lawrence V. Johnson Library was completed in 2015, but the libraries' physical collections have not been integrated
- KSU library Collection Development Department started systematic weeding project in 2017

# Weeding project

## Primary concerns:

- Age of the collection
- Irrelevance to the curriculum
- Item usage/lack of use
- Duplications among different campuses



# Data for the project -- Criteria

The Collection Development Department would like to have the data to review books published before 2008 that have not been checked out in the past 3 years, for the purpose of weeding



# Purpose of the Presentation

We would like to demonstrate

- ✓ How many reports we created
- ✓ What Criteria we used
- ✓ How we did the reports
- ✓ The problems we encountered
- ✓ How we solved our problems
- ✓ Remaining issues for further investigation



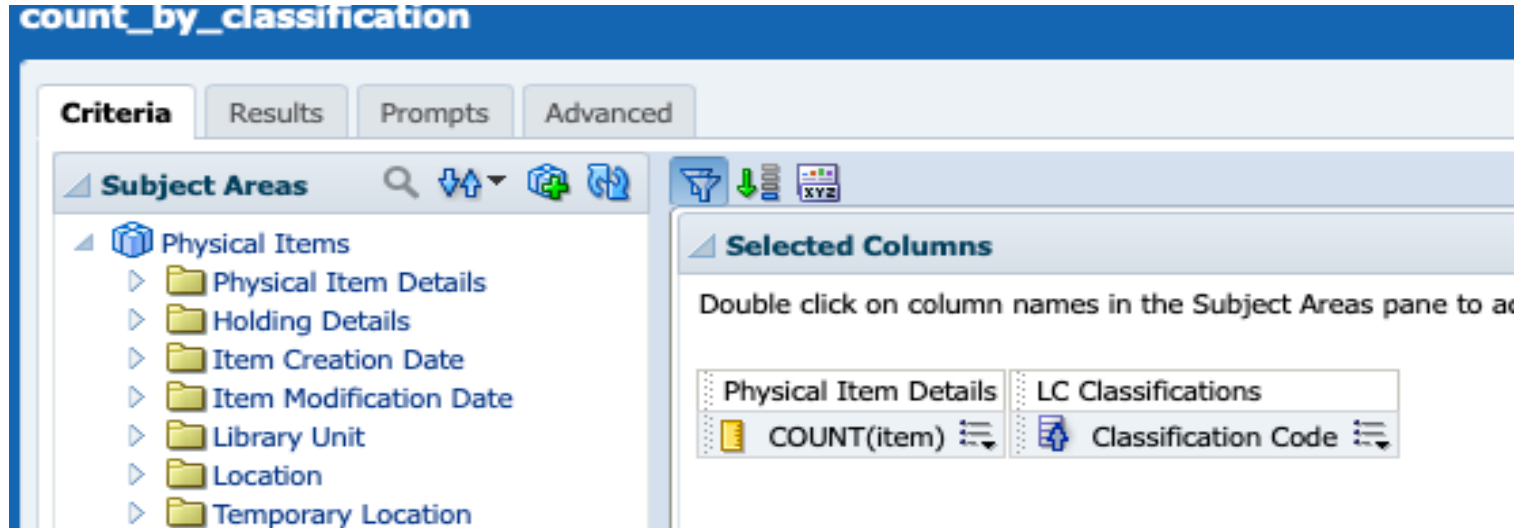
# Overview of Reports created

1. Count number of items in each classification
2. Shelflist published including 2008 and before with classification code
3. Shelflist published including 2008 and before with call number
4. Shelflist published after 2008 with classification code
5. Shelflist published after 2008 with call number
6. Books checked out before 2008
7. Shelflist before 2008 excluding checkouts
8. E-book and physical inventory comparison



# Item Count Report

1. Count number of items in each classification
  - count function (barcode)
  - Classification code



# Item Count Report --Result

ExLibris Analytics

count\_by\_classification

Criteria Results Prompts Advanced

Subject Areas

- Physical Items
  - Physical Item Details
  - Holding Details
  - Item Creation Date
  - Item Modification Date
  - Library Unit
  - Location
  - Temporary Location
  - Bibliographic Details
  - LC Classifications
  - Dewey Classifications
  - PO Line
  - Fund Information
  - Institution

Catalog

Compound Layout

Title: count\_by\_classification

Table:

COUNT(item)	Classification Code
454	AC
340	AE
124	AG
268	AI
42	AM
551	AP
141	AS
55	AY
92	AZ
3,172	B
244	BC
602	BD
6,468	BF
150	BH
855	BJ
1,849	BL

# Shelflist reports

After discussion with the collection librarian, we decided to use following fields:

- Title
- Author
- Barcode
- Publisher
- Location Name
- Year of Publication (\*)
- Lifecycle (\*)
- Processing Type (\*)
- Call Number (\*)/ Classification code
- Edition (\*)
- Volume (\*)



# Important elements

- *Lifecycle* is equal to / is in Active  
(without this, queries will contain deleted items)
- *Process Type* is equal to None  
(to filter out loan, lost and missing books etc)
- Call number vs classification code



# Important elements

We realized there were records that had the same title and needed to differentiate them  
So we added

- Edition
- Volume



# E-books and Physical Inventory Comparison

- Only needed purchased e-books. E-books available through subscription were not considered
- If books were available as an e-book and a physical item, the physical book could be considered for weeding
- Report 1 created as a Funds Expenditure report to only show e-books that were purchased
- Report 2 compared Report 1 against the shelf list of books published before 2008



# Example – E-books comparison

**Selected Columns**

Double click on column names in the Subject Areas pane to add them to the analysis. Once added, drag-and-drop columns to reorder them. Edit a column's properties, formula and filters, apply sorting, or delete by clicking on the column header.

Physical Item Details		Location		Bibliographic Details		Holding Details		Physical Item Details		Bibliographic Details	
Barcode	Library Name	Location Name	Title	Author	Permanent Call Number	Item Copy Id	Description	Publication Date	Publisher		

**Filters**

Add filters to the analysis criteria by clicking on Filter option for the specific column in the Selected Columns pane, or by clicking on the filter button in the Filter pane header. Add a saved filter by clicking on add button after the filter is created.

- Library Name is equal to / is in Johnson Library; Repository; Sturgis Library
- AND** Location Name is equal to / is in J-General; J-Oversize; R-General; S-General
- AND** Holding Lifecycle is equal to / is in Active
- AND** Material Type is equal to / is in Book
- AND** Classification Code is equal to / is in T; TA; TC; TD; TE; TF; TG; TH; TJ; TK; TL; TN; TP; TR; TS; TT; TX;
- AND** Publication Date is not LIKE (pattern match) %2008%; %2009%; %2010%; %2011%; %2012%; %2013%; %2014%; %2015%; %2016%; %2017%; %2018%; %2019%
- AND** Title (Normalized) is equal to any Title (Normalized) in Ebook2



# Challenge one

Year of Publication format:

2008	[2008 printing]	2007 [i.e. 2008]	[2008]-
c2008	©1982-©2008.	2008 [i.e. 2007]	a2008.
[2008]	2008-[2012]	[2010], c2008.	c2000-c2008.
[c2008]	Copyright 2008	[2008]-	C2008.;
©2008.	c [2008]	[©2008], ®1993.	c2008-



## Challenge one – cont2.

Because characters were alphanumeric and there was a lack of standard format, we could not simply use less than/greater than as a filter



# Challenge one – solution

We used wildcard character with "IS LIKE" and "Is NOT LIKE" operator

% = multiple characters (\*)

\_ = single character

Filter out

%2008%; %2009%; %2010%; %2011%;

%2012%; %2013%; %2014%; %2015%;

%2016%; %2017%; %2018%; %2019%;



# Challenge Two

## Call Number:

- Classification Code (LC Classifications)
- Permanent Call Number (Holding Details)

- Wildcard need to be used

- Example for Q and QA

Q1%;Q2%;Q3%...Q9%

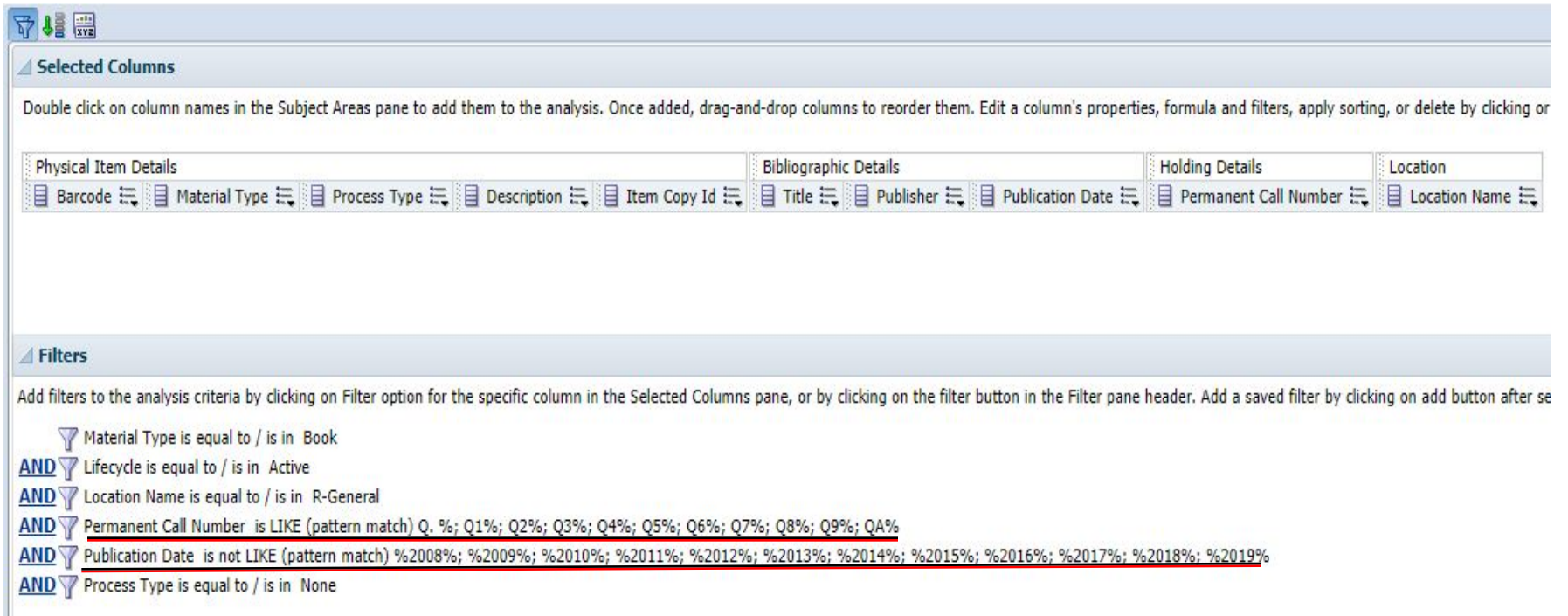
QA%

We avoid using Q% on this case, QC-QX call number will not be included



**KENNESAW STATE**  
UNIVERSITY  
LIBRARY SYSTEM

# Example: Shelf list report



**Selected Columns**

Double click on column names in the Subject Areas pane to add them to the analysis. Once added, drag-and-drop columns to reorder them. Edit a column's properties, formula and filters, apply sorting, or delete by clicking on the delete icon.

Physical Item Details	Bibliographic Details	Holding Details	Location
Barcode	Title	Permanent Call Number	Location Name
Material Type	Publisher		
Process Type	Publication Date		
Description			
Item Copy Id			

**Filters**

Add filters to the analysis criteria by clicking on Filter option for the specific column in the Selected Columns pane, or by clicking on the filter button in the Filter pane header. Add a saved filter by clicking on add button after selection.

- Material Type is equal to / is in Book
- AND Lifecycle is equal to / is in Active
- AND Location Name is equal to / is in R-General
- AND Permanent Call Number is LIKE (pattern match) Q. %; Q1%; Q2%; Q3%; Q4%; Q5%; Q6%; Q7%; Q8%; Q9%; QA%
- AND Publication Date is not LIKE (pattern match) %2008%; %2009%; %2010%; %2011%; %2012%; %2013%; %2014%; %2015%; %2016%; %2017%; %2018%; %2019%
- AND Process Type is equal to / is in None



# Challenge Three

We needed checkout data from 2016 to current. We migrated to Alma in May 2017, and according to Alma to Voyager Migration Guide:

“Active loan transactions are migrated from Voyager to Alma.

Historical circulation activity in the form of counts (e.g. historical charges, last loan date, etc.)”



# Challenge Three -- solution

So we had to get reports from Access, then manually combined Voyager loans with Analytics loans in Excel sheet and took out the duplicated records



**KENNESAW STATE**  
UNIVERSITY  
LIBRARY SYSTEM

# Challenge Four

- We needed to exclude checkout data from our shelflist report because books being checked out were not being considered for weeding
- Using the checkout report, we excluded the checkout data from the shelflist report
- Some of the comparative reports in Alma Analytics returned errors, so we had to manually remove duplicated items by hand in Excel



# Example: Excluding checkout data report.

repository Q-QA shelflist before 2008 wildcard normal ...

Home Catalog Favorites Dashboards New Open Signed In A

Criteria Results Prompts Advanced

Subject Areas

- Physical Items
  - Physical Item Details
  - Holding Details
  - Item Creation Date
  - Item Modification Date
  - Library Unit
  - Location
  - Temporary Location
  - Bibliographic Details
  - LC Classifications
  - Dewey Classifications
  - PO Line
  - Fund Information
  - Institution

Selected Columns

Double click on column names in the Subject Areas pane to add them to the analysis. Once added, drag-and-drop columns to reorder them. Edit a column's properties, formula and filters, apply sorting, or delete by clicking or hovering over the button next to its name.

Physical Item Details	Bibliographic Details	Holding Details	Location
Barcode	Title	Permanent Call Number	Location Name

Filters

Add filters to the analysis criteria by clicking on Filter option for the specific column in the Selected Columns pane, or by clicking on the filter button in the Filter pane header. Add a saved filter by clicking on add button after selecting its name in the catalog pane.

- Material Type is equal to / is in Book
- AND Lifecycle is equal to / is in Active
- AND Location Name is equal to / is in R-General
- AND Permanent Call Number is LIKE (pattern match) Q. %; Q1%; Q2%; Q3%; Q4%; Q5%; Q6%; Q7%; Q8%; Q9%; QA%
- AND Publication Date is not LIKE (pattern match) %2008%; %2009%; %2010%; %2011%; %2012%; %2013%; %2014%; %2015%; %2016%; %2017%; %2018%; %2019%
- AND Process Type is equal to / is in None
- AND Barcode is equal to any Barcode in repository Q-QA before 2008 checkout wildcard

# Example: Removing data by hand

The screenshot shows the Microsoft Excel interface with the 'Conditional Formatting' menu open. The menu options include: Highlight Cells Rules, Top/Bottom Rules, Data Bars, Color Scales, Icon Sets, New Rule..., Clear Rules, and Manage Rules... The 'Highlight Cells Rules' sub-menu is also visible, showing options like Greater Than..., Less Than..., Between..., Equal To..., Text that Contains..., A Date Occurring..., and Duplicate Values... The spreadsheet data is as follows:

	A	B
1	T-TX shelflist before 2008 permanent call number - Jason NEW TEST	
2		
4		Mechanics for engineers--statics and dynamics /
5		Statics and strength of materials /
6		Systematic layout planning /
255	30427000231968	Principles of construction management /
	30427000232420	Energy-effective windows : the search for knowledge engineering of windows in the cause of energy conservation : joint DOE (ERDA)/NBS conference/round table on energy conservation in Washington, D.C., on April 13th, 1977. /
283	30427000234301	The Encyclopedia of how it's built /
320	30427000234434	Building in England down to 1540: a documentary history, /
329	30427000234541	The building of Manhattan /
338	30427000234764	The use of CPM in construction : a manual for general contractors and the construction industry /
356		
20246		
20263	30427000231968	
20264	30427000232420	
20265	30427000234301	
20266	30427000234434	
20267	30427000234541	
20268	30427000234764	



# Challenge Five


- When we started creating reports in Analytics, there was a field called “Times Loaned - not sum” in Physical Items subject. We used it to get checkout counts
- After generating the report, we took a sample of data, examined it in Alma Fulfilment, and found a lot of discrepancy
- Then we found following explanation from Alma site:

# Challenge Five – Exlibris explanaton

Search under current path



Alma

 ▾ | Home » Alma » Knowledge Articles » What is the difference between "Time Loaned" and "Times Loaned - not sum"

## What is the difference between "Time Loaned" and "Times Loaned - not sum"



- **Article Type:** Q&A
- **Product:** Alma

### Question

In Alma Analytics, in the **Physical Items subject** area there are two fields "Time Loaned" and "Times Loaned - not sum". What is the difference between them?

### Answer

Please use the "Time Loaned" field only. The "Times Loaned - not sum" field does not deliver reliable results and will probably be removed in a future release.

**Category:** Alma Analytics - Alma

- **Article last edited:** 6/26/2014



**KENNESAW STATE**  
**UNIVERSITY**  
LIBRARY SYSTEM

# Conclusion

1. Count number of items in each classification
2. Shelflist published before 2008 excluding checkout data with classification code
3. Shelflist published excluding before 2008 with call number
4. Shelflist published in and after 2008 with classification code
5. Shelflist published in and after 2008 with call number
6. Books checked out before 2008
7. E-book and physical inventory comparison



# Issues need to be solved in the future

- The last copy issue
- Compare our collection to other USG institutions using NZ report
- Compare our collection to HathiTrust report



# Another resource and a tip

- Ex Libris Report: Physical Usage for Weeding
- Remember to take a small sampling of data to check in Alma to validate the report data



Thanks for your attention..

Any questions?



**KENNESAW STATE**  
UNIVERSITY  
LIBRARY SYSTEM