2016-10-14

Using OpenRefine for data-driven decision-making: a case study from the University of Michigan Library

Carruthers, Matthew

https://hdl.handle.net/2027.42/138109

http://creativecommons.org/licenses/by-nc-sa/4.0/

Downloaded from Deep Blue, University of Michigan's institutional repository

USING OPENREFINE FOR DATA-DRIVEN DECISION-MAKING

A case study from the University of Michigan Library

Matt Carruthers
Metadata Projects Librarian
University of Michigan

OpenRefine



openrefine.org

OpenRefine



openrefine.org

(He's lying.)

Project Summary

- Identify resources in our collection for which we own more than one copy.
- Assess whether or not they are low-use items.
- Withdraw duplicate copies from the collection or send to remote storage.

Project Summary

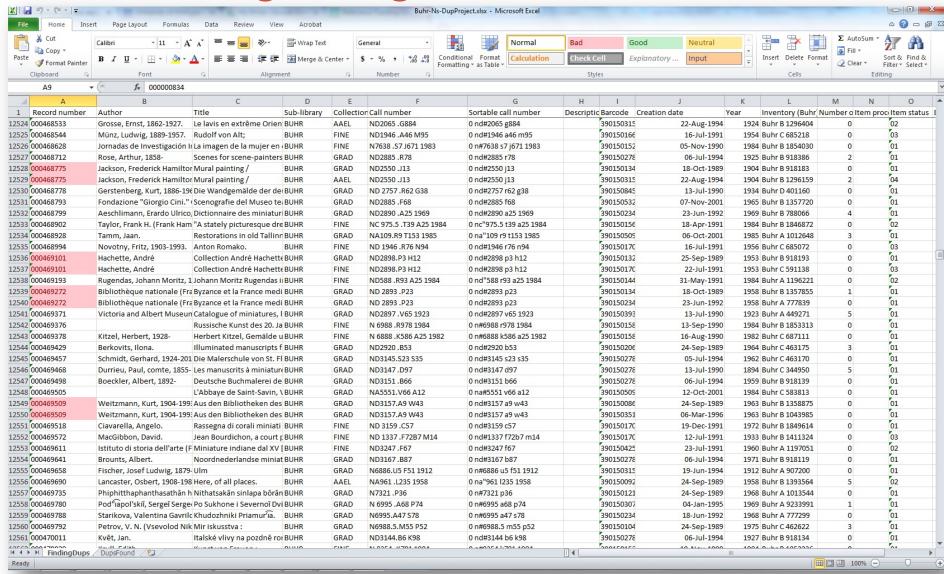
DataMart

- Library's reporting tool for extracting data from our ILS.
- Reports are generated in spreadsheet format.
- Spreadsheets contain a mix of bibliographic information and usage statistics.

Project Summary

- Find items which:
 - Have more than one copy in the Library.
 - Are not part of a series.
 - Are at least five years old.
 - Have circulated less than six times.
 - Are not "in process".
 - Are not attached to a provisional cataloging record.
 - Are part of the circulating collection.

In the Beginning



Google refine MMDP test run Permalink

Open... Export ▼ Help

Facet / Filter Undo / Redo 2

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

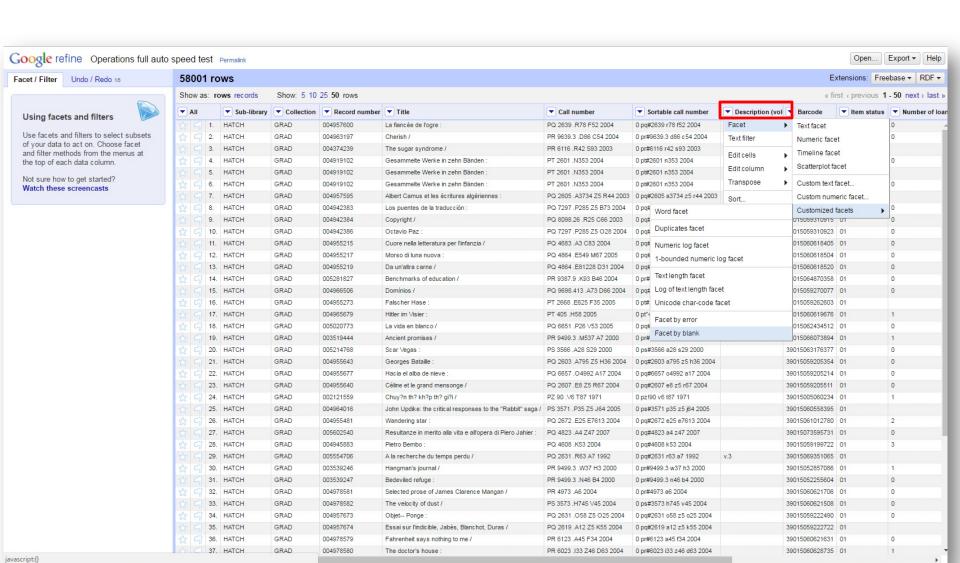
Not sure how to get started? Watch these screencasts

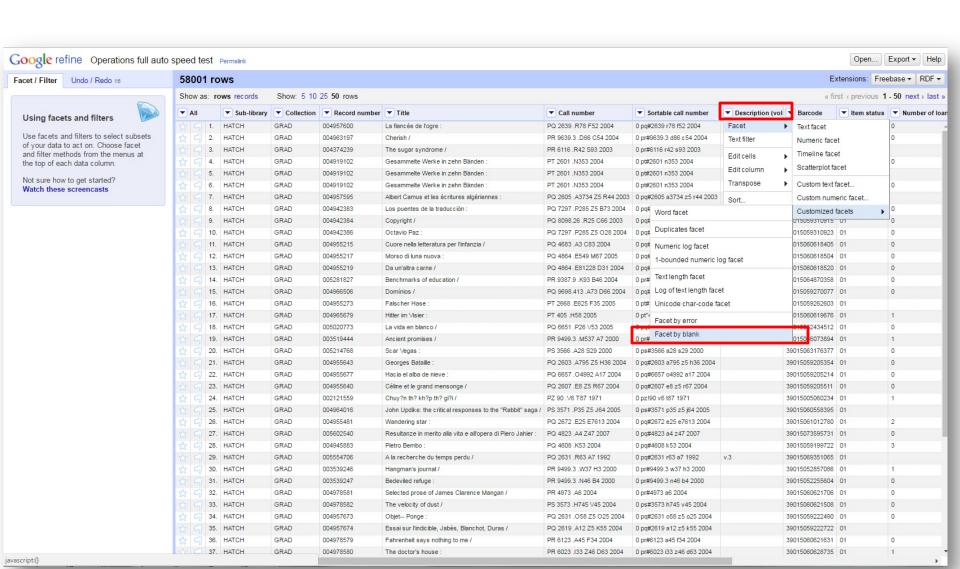
99992 rows

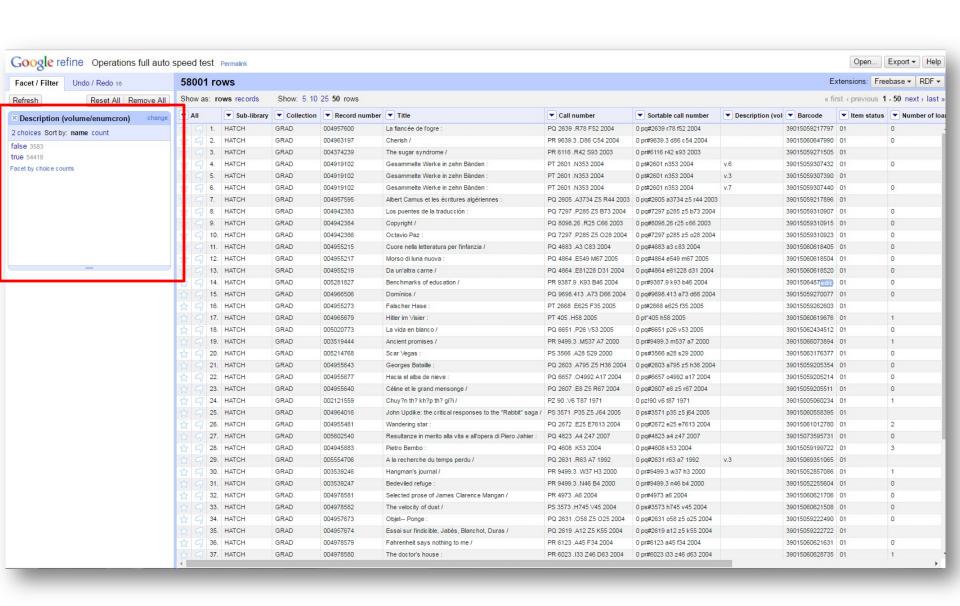
Show as: rows records Show: 5 10 25 50 rows

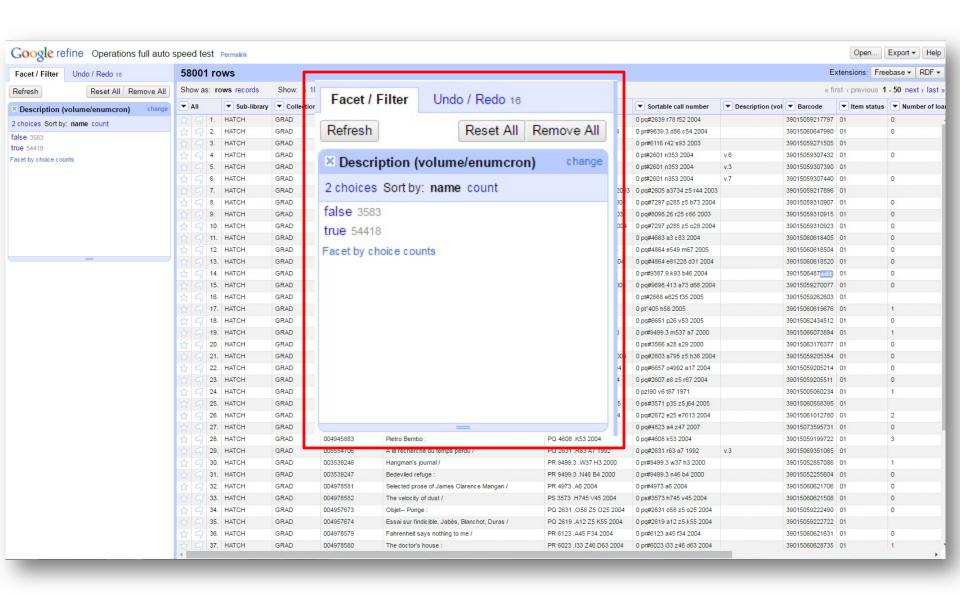
Extensions: Freebase ▼ RDF ▼ « first ¿ previous 1 - 50 next » last »

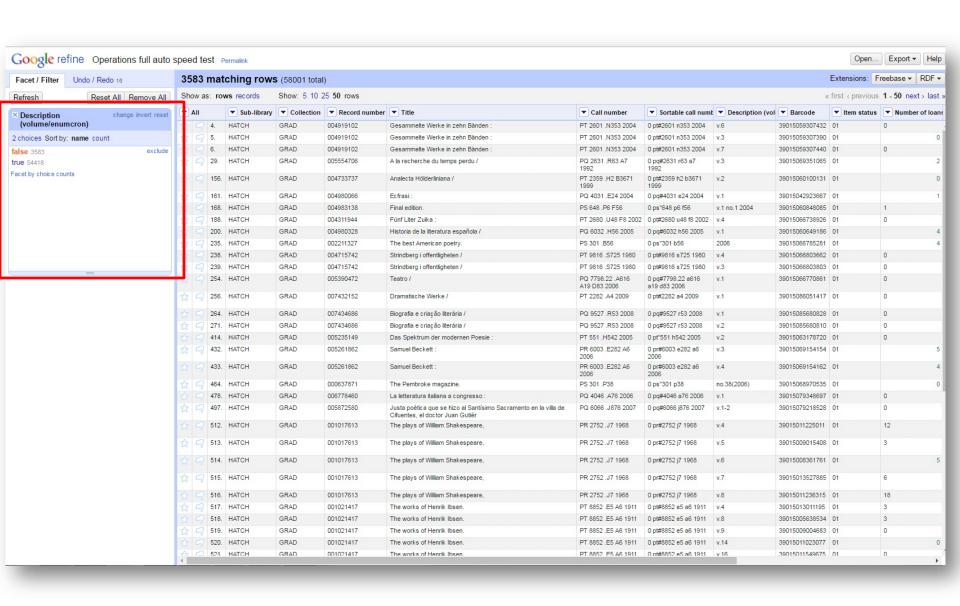
AI	II		▼ Sub-library	Collection	Record number	▼ Title	▼ Call number	▼ Sortable call number	Description (volume/en	▼ Barcode	▼ Item status	Number of loans	•
		1.	HATCH	GRAD	004957600	La fiancée de l'ogre :	PQ 2639 .R78 F52 2004	0 pq#2639 r78 f52 2004		39015059217797	01	0	21 🗚
		2.	HATCH	GRAD	004963197	Cherish /	PR 9639.3 .D86 C54 2004	0 pr#9639.3 d86 c54 2004		39015060647990	01	0	21
		3.	HATCH	GRAD	011717137	Le dernier amour d'Arsène Lupin :	PQ 2623 .E145 D47 2012	0 pq#2623 e145 d47 2012		39015091350267	01	0	21
		4.	HATCH	GRAD	011717144	L'âme du monde /	PQ 2672 .E5642 A67 2012	0 pq#2672 e5642 a67 2012		39015091350275	01	0	21
		5.	HATCH	GRAD	011717152	Henri Meschonnic dans tous ses états :	PQ 2673 .E712 Z73 2012	0 pq#2673 e712 z73 2012		39015091350317	01	0	21
		6.	HATCH	GRAD	011660612	Melyador :	PQ 1461 .F8 M34 2012	0 pq#1461 f8 m34 2012	v.2	39015091204837	01	0	21
		7.	HATCH	GRAD	004374239	The sugar syndrome /	PR 6116 .R42 S93 2003	0 pr#6116 r42 s93 2003		39015059271505	01	0	21
		8.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden :	PT 2601 .N353 2004	0 pt#2601 n353 2004	v.6	39015059307432	01	0	21
		9.	HATCH	GRAD	011660767	Vivir sin los otros :	PQ 8180.417 .O593 V58 2010	0 pq#8180.417 o593 v58 2010		39015091345853	01	0	21
		10.	HATCH	GRAD	012234286	Voices in the wind :	PR 9402.5 .V65 2010	0 pr#9402.5 v65 2010		39015092024176	01	0	21
		11.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden :	PT 2601 .N353 2004	0 pt#2601 n353 2004	v.3	39015059307390	01	0	21
		12.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden :	PT 2601 .N353 2004	0 pt#2601 n353 2004	v.7	39015059307440	01	0	21
		13.	HATCH	GRAD	004957595	Albert Camus et les écritures algériennes :	PQ 2605 .A3734 Z5 R44 2003	0 pq#2605 a3734 z5 r44 2003		39015059217896	01	3	21
		14.	HATCH	GRAD	004942383	Los puentes de la traducción :	PQ 7297 .P285 Z5 B73 edit 2004	0 pq#7297 p285 z5 b73 2004		39015059310907	01	0	21
		15.	HATCH	GRAD	004942384	Copyright /	PQ 8098.26 .R25 C66 2003	0 pq#8098.26 r25 c66 2003		39015059310915	01	0	21
		16.	HATCH	GRAD	004942386	Octavio Paz :	PQ 7297 .P285 Z5 O28 2004	0 pq#7297 p285 z5 o28 2004		39015059310923	01	0	21
		17.	HATCH	GRAD	012228383	Reflections & retrospectives /	PR 9344 .R44 2012	0 pr#9344 r44 2012		39015091965817	01	1	21
		18.	HATCH	GRAD	012229441	L'ombre de mémoire /	PQ 2605 .O33 O43 1990	0 pq#2605 o33 o43 1990		39015091372972	01	0	1!
		19.	HATCH	GRAD	004955215	Cuore nella letteratura per l'infanzia /	PQ 4683 .A3 C83 2004	0 pq#4683 a3 c83 2004		39015060618405	01	0	21
		20.	HATCH	GRAD	004955217	Morso di luna nuova :	PQ 4864 .E549 M67 2005	0 pq#4864 e549 m67 2005		39015060618504	01	0	21
		21.	HATCH	GRAD	004955219	Da un'altra carne /	PQ 4864 .E81228 D31 2004	0 pq#4864 e81228 d31 2004		39015060618520	01	0	21
		22.	HATCH	GRAD	012231343	Pour l'histoire du "Thésée" d'André Gide :	PQ 2613 .I2 T52 2012	0 pq#2613 i2 t52 2012		39015091930571	01	0	21
		23.	HATCH	GRAD	005281827	Benchmarks of education /	PR 9387.9 .K93 B46 2004	0 pr#9387.9 k93 b46 2004		39015064870358	01	0	21
		24.	HATCH	GRAD	004966506	Domínios /	PQ 9698.413 .A73 D66 2004	0 pq#9698.413 a73 d66 2004		39015059270077	01	0	21
		25.	HATCH	GRAD	004955273	Falscher Hase :	PT 2668 .E625 F35 2005	0 pt#2668 e625 f35 2005		39015059262603	01	1	21
		26.	HATCH	GRAD	000677633	Poetry.	PS 301 .P74	0 ps"301 p74	v.200:no.1-2,4(2012:Apr Aug.)	39015091146822	01	0	1!
		27.	HATCH	GRAD	012230958	Black box played /	PR 9381.9 .O2 B53 2010	0 pr#9381.9 o2 b53 2010		39015091978612	01	0	21
		28.	HATCH	GRAD	004965679	Hitler im ∀isier :	PT 405 .H58 2005	0 pt"405 h58 2005		39015060619676	01	1	21
		29.	HATCH	GRAD	005020773	La vida en blanco /	PQ 6651 .P26 V53 2005	0 pq#6651 p26 v53 2005		39015062434512	01	0	21
		30.	HATCH	GRAD	011688259	Goodbye, Flicker:	PS 3607 .145215 M36 2012	0 ps#3607 i45215 m36 2012		39015088282234	01	1	21
		31.	HATCH	GRAD	012234274	Africa's new brood :	PR 9402.9 .W364 A69 2006	0 pr#9402.9 w364 a69 2006		39015092023871	01	0	21
		32.	HATCH	GRAD	003519444	Ancient promises /	PR 9499.3 .M537 A7 2000	0 pr#9499.3 m537 a7 2000		39015066073894	01	1	21
		33.	HATCH	GRAD	012238561	Never too late /	PR 9402.8 .N48 2010	0 pr#9402.8 n48 2010		39015092023632	01	0	21
5 1		3.4	нлтон	CRAD	D12238564	Namihian youth theatre	DE CAUS NISSO NISS OUUS	0 nr#0/08 n362 n36 2008		30015002023840	Π1	n	21











Extensions: Freebase ▼ RDF ▼

« first « previous 1 - 50 next » last »

Facet / Filter Undo / Redo 2

Using facets and filters

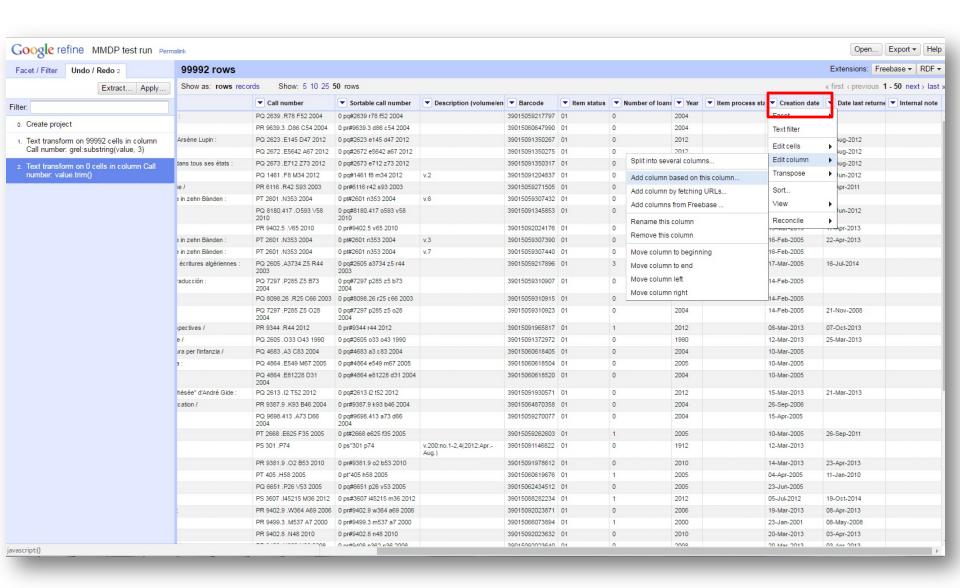
Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

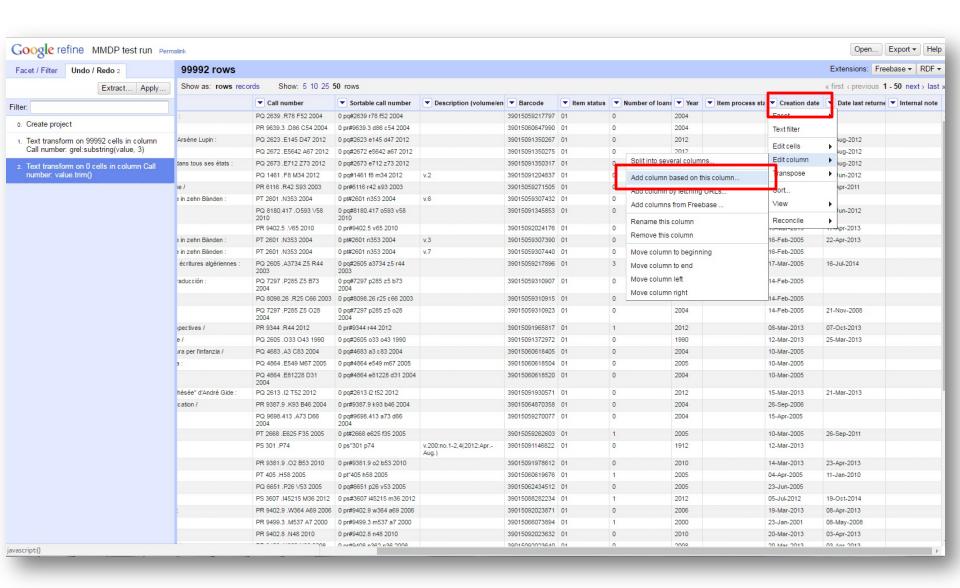
Not sure how to get started? Watch these screencasts

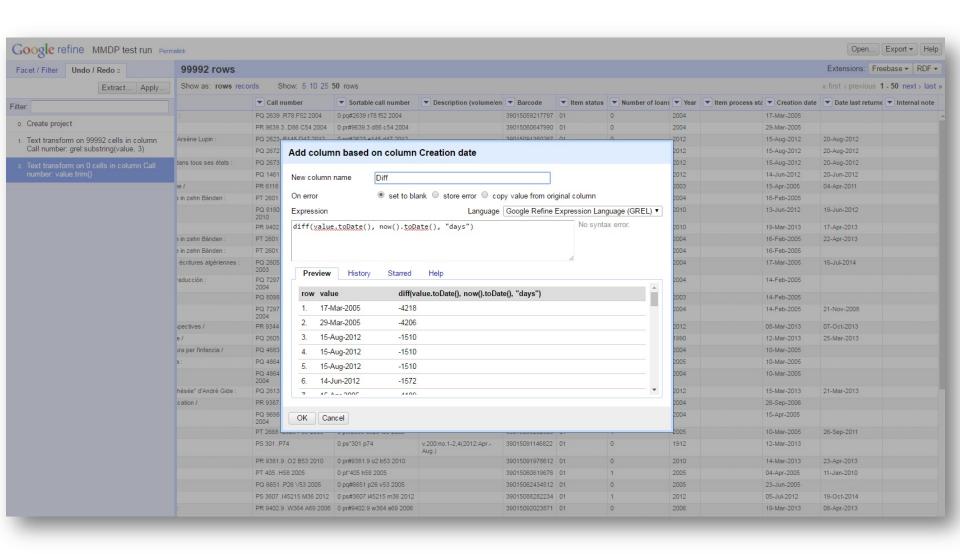
99992 rows

Show as: rows records Show: 5 10 25 50 rows

Title	▼ Call number	▼ Sortable call number	▼ Description (volume/en	▼ Barcode	▼ Item status	▼ Number of loans	▼ Year	▼ Item process sta	Creation date	Date last returne
iancée de l'ogre :	PQ 2639 .R78 F52 2004	0 pq#2639 r78 f52 2004		39015059217797	01	0	2004		17 Mar 2005	
rish /	PR 9639.3 .D86 C54 2004	0 pr#9639.3 d86 c54 2004		39015060647990	01	0	2004		29-Mar-2005	
dernier amour d'Arsène Lupin :	PQ 2623 .E145 D47 2012	0 pq#2623 e145 d47 2012		39015091350267	01	0	2012		15-Aug-2012	20-Aug-2012
ne du monde /	PQ 2672 .E5642 A67 2012	0 pq#2672 e5642 a67 2012		39015091350275	01	0	2012		15-Aug-2012	20-Aug-2012
ri Meschonnic dans tous ses états :	PQ 2673 .E712 Z73 2012	0 pq#2673 e712 z73 2012		39015091350317	01	0	2012		15-Aug-2012	20-Aug-2012
yador :	PQ 1461 .F8 M34 2012	0 pq#1461 f8 m34 2012	v.2	39015091204837	01	0	2012		14-Jun-2012	20-Jun-2012
sugar syndrome /	PR 6116 .R42 S93 2003	0 pr#6116 r42 s93 2003		39015059271505	01	0	2003		15-Apr-2005	04-Apr-2011
ammelte Werke in zehn Bänden :	PT 2601 .N353 2004	0 pt#2601 n353 2004	v.6	39015059307432	01	0	2004		16-Feb-2005	
r sin los otros :	PQ 8180.417 .O593 V58 2010	0 pq#8180.417 o593 v58 2010		39015091345853	01	0	2010		13-Jun-2012	19-Jun-2012
es in the wind :	PR 9402.5 .V65 2010	0 pr#9402.5 v65 2010		39015092024176	01	0	2010		19-Mar-2013	17-Apr-2013
ammelte Werke in zehn Bänden :	PT 2601 .N353 2004	0 pt#2601 n353 2004	v.3	39015059307390	01	0	2004		16-Feb-2005	22-Apr-2013
ammelte Werke in zehn Bänden :	PT 2601 .N353 2004	0 pt#2601 n353 2004	v.7	39015059307440	01	0	2004		16-Feb-2005	
rt Camus et les écritures algériennes :	PQ 2605 .A3734 Z5 R44 2003	0 pq#2605 a3734 z5 r44 2003		39015059217896	01	3	2004		17-Mar-2005	16-Jul-2014
puentes de la traducción :	PQ 7297 .P285 Z5 B73 2004	0 pg#7297 p285 z5 b73 2004		39015059310907	01	0	2004		14-Feb-2005	
right /	PQ 8098.26 .R25 C66 2003	0 pq#8098.26 r25 c66 2003		39015059310915	01	0	2003		14-Feb-2005	
avio Paz :	PQ 7297 .P285 Z5 O28 2004	0 pg#7297 p285 z5 o28 2004		39015059310923	01	0	2004		14-Feb-2005	21-Nov-2008
ections & retrospectives /	PR 9344 .R44 2012	0 pr#9344 r44 2012		39015091965817	01	1	2012		08-Mar-2013	07-Oct-2013
bre de mémoire /	PQ 2605 .O33 O43 1990	0 pq#2605 o33 o43 1990		39015091372972	01	0	1990		12-Mar-2013	25-Mar-2013
re nella letteratura per l'infanzia /	PQ 4683 .A3 C83 2004	0 pq#4683 a3 c83 2004		39015060618405	01	0	2004		10-Mar-2005	
so di luna nuova :	PQ 4864 .E549 M67 2005	0 pq#4864 e549 m67 2005		39015060618504	01	0	2005		10-Mar-2005	
ın'altra carne /	PQ 4864 .E81228 D31 2004	0 pq#4864 e81228 d31 2004		39015060618520	01	0	2004		10-Mar-2005	
r l'histoire du "Thésée" d'André Gide :	PQ 2613 .12 T52 2012	0 pq#2613 i2 t52 2012		39015091930571	01	0	2012		15-Mar-2013	21-Mar-2013
chmarks of education /	PR 9387.9 .K93 B46 2004	0 pr#9387.9 k93 b46 2004		39015064870358	01	0	2004		26-Sep-2006	
nínios /	PQ 9698.413 .A73 D66 2004	0 pq#9698.413 a73 d66 2004		39015059270077	01	0	2004		15-Apr-2005	
cher Hase :	PT 2668 .E625 F35 2005	0 pt#2668 e625 f35 2005		39015059262603	01	1	2005		10-Mar-2005	26-Sep-2011
try.	PS 301 .P74	0 ps"301 p74	v.200:no.1-2,4(2012:Apr Aug.)	39015091146822	01	0	1912		12-Mar-2013	
k box played /	PR 9381.9 .O2 B53 2010	0 pr#9381.9 o2 b53 2010		39015091978612	01	0	2010		14-Mar-2013	23-Apr-2013
rim Visier:	PT 405 .H58 2005	0 pt"405 h58 2005		39015060619676	01	1	2005		04-Apr-2005	11-Jan-2010
ida en blanco /	PQ 6651 .P26 V53 2005	0 pq#6651 p26 v53 2005		39015062434512	01	0	2005		23-Jun-2005	
dbye, Flicker:	PS 3607 .145215 M36 2012	0 ps#3607 i45215 m36 2012		39015088282234	01	1	2012		05-Jul-2012	19-Oct-2014
a's new brood :	PR 9402.9 .W364 A69 2006	0 pr#9402.9 w364 a69 2006		39015092023871	01	0	2006		19-Mar-2013	08-Apr-2013
ient promises /	PR 9499.3 .M537 A7 2000	0 pr#9499.3 m537 a7 2000		39015066073894	01	1	2000		23-Jan-2001	08-May-2008
er too late /	PR 9402.8 .N48 2010	0 pr#9402.8 n48 2010		39015092023632	01	0	2010		20-Mar-2013	03-Apr-2013
nihian vouth theatre	DR GADS NISES NISE SOOR	0 nr#0408 n362 n36 2008		20015002023840	Λ1	n	2008		20 Mar 2013	03 Apr 2013





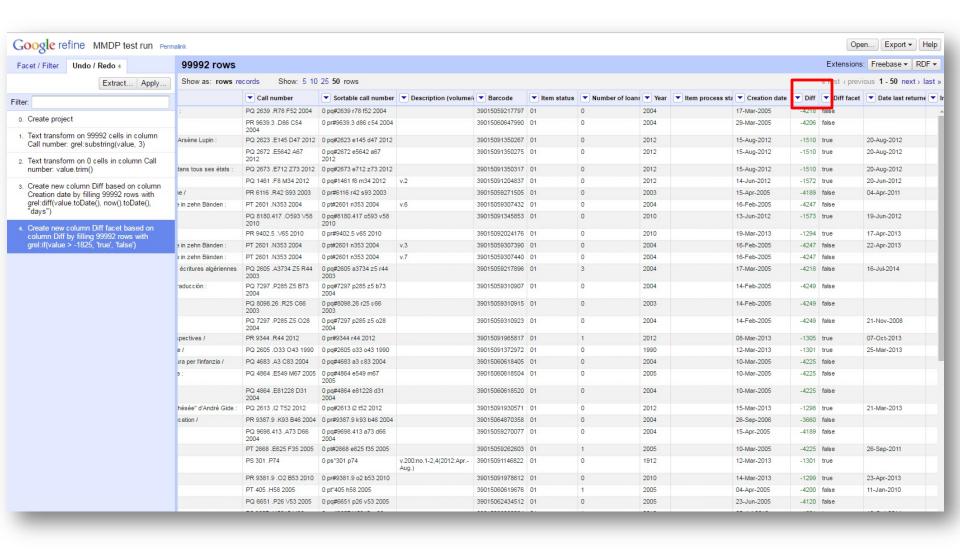


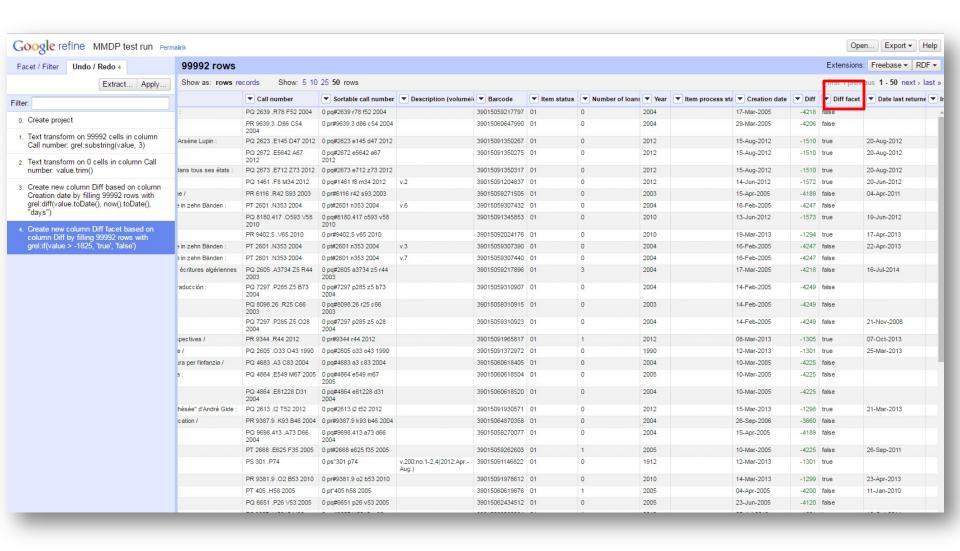
Add column based on column Creation date

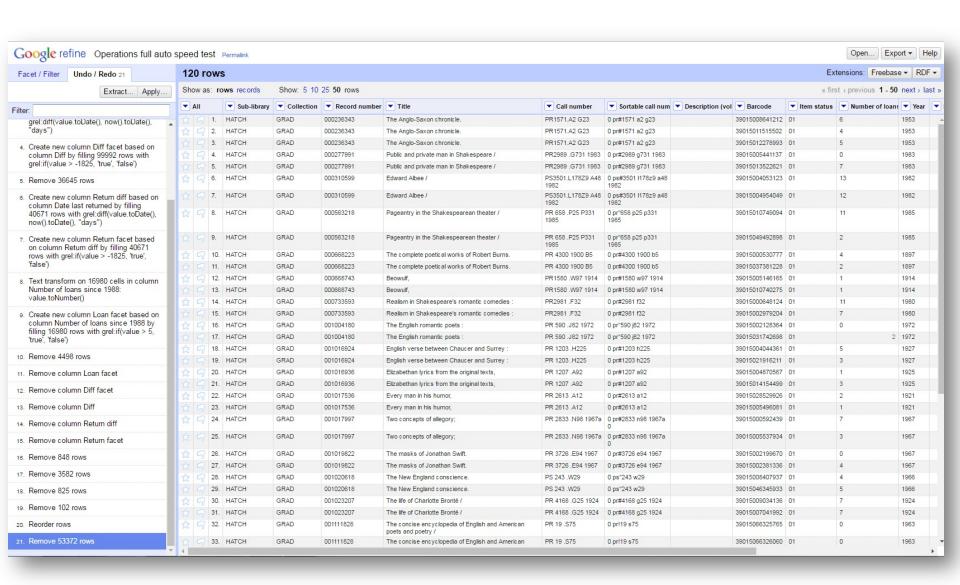
New column name	Diff	
On error	set to blank ○ store error ○ copy value.	ue from original column
Expression	Language Goog	gle Refine Expression Language (GREL) ▼
diff(value.toDate	(), now().toDate(), "days")	No syntax error.

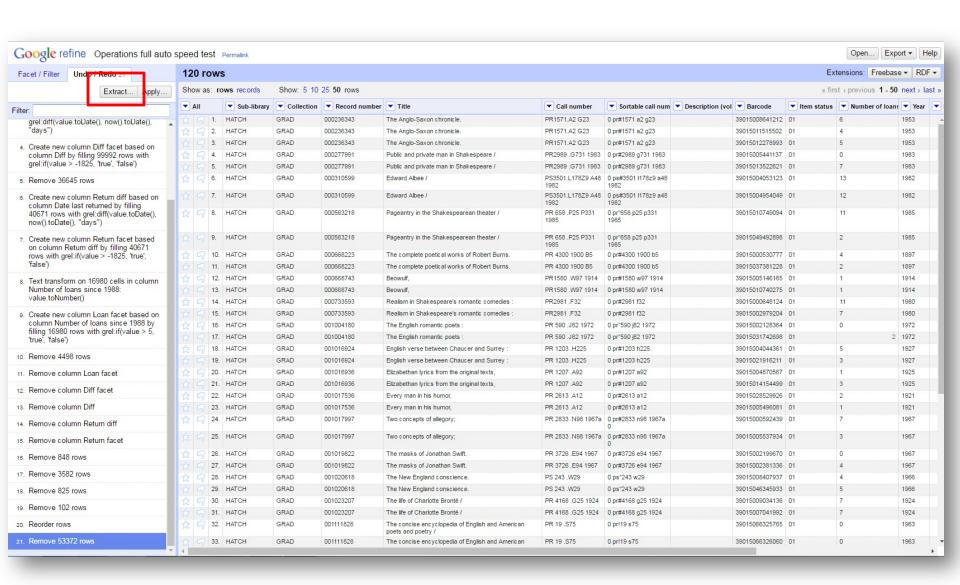
row	value	diff(value.toDate(), now().toDate(), "days")	1
1.	17-Mar-2005	-4218	
2.	29-Mar-2005	-4206	
3.	15-Aug-2012	-1510	
4.	15-Aug-2012	-1510	
5.	15-Aug-2012	-1510	
6.	14-Jun-2012	-1572	
7	4F A 200F	4400	

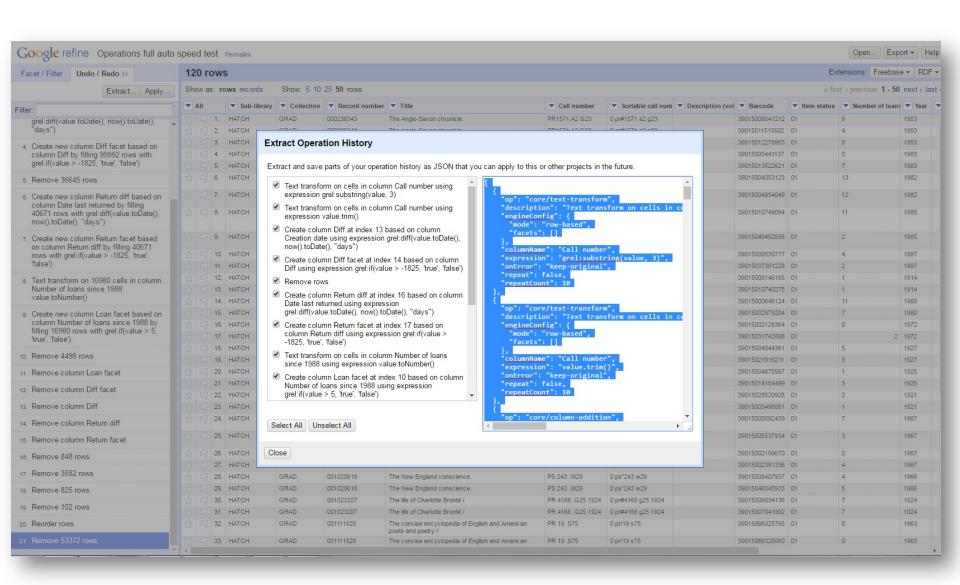
OK Cancel











Google refine PQ PZGrad csv Permalink

Open... Export ▼ Help

Facet / Filter Undo / Redo o



Infinite undo history

Don't worry about making mistakes. Every change you make will be shown here, and you can undo your changes anytime.

Learn more »



99992 rows Extensions: Freebase ▼ RDF ▼ Show as: rows records Show: 5 10 25 50 rows « first « previous 1 - 50 next » last »

V	All Sub-library		Collection	Record number	▼ Title	▼ Call number	 Sortable call number 	▼ Description (volume/en	▼ Barcode	▼ Item status	▼ Number of loans	
		1.	HATCH	GRAD	004957600	La fiancée de l'ogre :	\$\$hPQ 2639 .R78 F52 2004	0 pq#2639 r78 f52 2004		39015059217797	01	0
		2.	HATCH	GRAD	004963197	Cherish /	\$\$hPR 9639.3 .D86 C54 2004	0 pr#9639.3 d86 c54 2004		39015060647990	01	0
		3.	HATCH	GRAD	011717137	Le dernier amour d'Arsène Lupin :	\$\$hPQ 2623 .E145 D47 2012	0 pq#2623 e145 d47 2012		39015091350267	01	0
		4.	HATCH	GRAD	011717144	L'âme du monde /	\$\$hPQ 2672 .E5642 A67 2012	0 pq#2672 e5642 a67 2012		39015091350275	01	0
		5.	HATCH	GRAD	011717152	Henri Meschonnic dans tous ses états :	\$\$hPQ 2673 .E712 Z73 2012	0 pq#2673 e712 z73 2012		39015091350317	01	0
		6.	HATCH	GRAD	011660612	Melyador :	\$\$hPQ 1461 .F8 M34 2012	0 pq#1461 f8 m34 2012	v.2	39015091204837	01	0
		7.	HATCH	GRAD	004374239	The sugar syndrome /	\$\$hPR 6116 .R42 S93 2003	0 pr#6116 r42 s93 2003		39015059271505	01	0
		8.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden :	\$\$hPT 2601 .N353 2004	0 pt#2601 n353 2004	v.6	39015059307432	01	0
		9.	HATCH	GRAD	011660767	Vivir sin los otros :	\$\$hPQ 8180.417 .0593 \/58 2010	0 pq#8180.417 o593 v58 2010		39015091345853	01	0
		10	HATCH	GRAD	012234286	Voices in the wind :	\$\$hPR 9402.5 .V65 2010	0 pr#9402.5 v65 2010		39015092024176	01	0
		11.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden :	\$\$hPT 2601 .N353 2004	0 pt#2601 n353 2004	v.3	39015059307390	01	0
		12	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden :	\$\$hPT 2601 .N353 2004	0 pt#2601 n353 2004	v.7	39015059307440	01	0
		13	HATCH	GRAD	004957595	Albert Camus et les écritures algériennes :	\$\$hPQ 2605 .A3734 Z5 R44 2003	0 pq#2605 a3734 z5 r44 2003		39015059217896	01	3
		14	HATCH	GRAD	004942383	Los puentes de la traducción :	\$\$hPQ 7297 .P285 Z5 B73 2004	0 pq#7297 p285 z5 b73 2004		39015059310907	01	0
		15	HATCH	GRAD	004942384	Copyright /	\$\$hPQ 8098.26 .R25 C66 2003	0 pq#8098.26 r25 c66 2003		39015059310915	01	0
		16	HATCH	GRAD	004942386	Octavio Paz :	\$\$hPQ 7297 .P285 Z5 O28 2004	0 pq#7297 p285 z5 o28 2004		39015059310923	01	0
		17.	HATCH	GRAD	012228383	Reflections & retrospectives /	\$\$hPR 9344 .R44 2012	0 pr#9344 r44 2012		39015091965817	01	1
		18	HATCH	GRAD	012229441	L'ombre de mémoire /	\$\$hPQ 2605 .O33 O43 1990	0 pq#2605 o33 o43 1990		39015091372972	01	0
		19	HATCH	GRAD	004955215	Cuore nella letteratura per l'infanzia /	\$\$hPQ 4683 .A3 C83 2004	0 pq#4683 a3 c83 2004		39015060618405	01	0
		20	HATCH	GRAD	004955217	Morso di luna nuova :	\$\$hPQ 4864 .E549 M67 2005	0 pq#4864 e549 m67 2005		39015060618504	01	0
		21	HATCH	GRAD	004955219	Da un'altra carne /	\$\$hPQ 4864 .E81228 D31 2004	0 pq#4864 e81228 d31 2004		39015060618520	01	0
		22	HATCH	GRAD	012231343	Pour l'histoire du "Thésée" d'André Gide :	\$\$hPQ 2613 .I2 T52 2012	0 pq#2613 i2 t52 2012		39015091930571	01	0
		23	HATCH	GRAD	005281827	Benchmarks of education /	\$\$hPR 9387.9 .K93 B46 2004	0 pr#9387.9 k93 b46 2004		39015064870358	01	0
		24	HATCH	GRAD	004966506	Domínios /	\$\$hPQ 9698.413 .A73 D66 2004	0 pq#9698.413 a73 d66 2004		39015059270077	01	0
		25	HATCH	GRAD	004955273	Falscher Hase :	\$\$hPT 2668 .E625 F35 2005	0 pt#2668 e625 f35 2005		39015059262603	01	1
		26	HATCH	GRAD	000677633	Poetry.	\$\$hPS 301 .P74	0 ps"301 p74	v.200:no.1-2,4(2012:Apr Aug.)	39015091146822	01	0
		27	HATCH	GRAD	012230958	Black box played /	\$\$hPR 9381.9 .O2 B53 2010	0 pr#9381.9 o2 b53 2010		39015091978612	01	0
		28	HATCH	GRAD	004965679	Hitler im ∀isier :	\$\$hPT 405 .H58 2005	0 pt"405 h58 2005		39015060619676	01	1
		29	HATCH	GRAD	005020773	La vida en blanco /	\$\$hPQ 6651 .P26 V53 2005	0 pq#6651 p26 v53 2005		39015062434512	01	0
		30	HATCH	GRAD	011688259	Goodbye, Flicker :	\$\$hPS 3607 .145215 M36 2012	0 ps#3607 i45215 m36 2012		39015088282234	01	1
		31	HATCH	GRAD	012234274	Africa's new brood :	\$\$hPR 9402.9 .W364 A69 2006	0 pr#9402.9 w364 a69 2006		39015092023871	01	0
		32	HATCH	GRAD	003519444	Ancient promises /	\$\$hPR 9499.3 .M537 A7 2000	0 pr#9499.3 m537 a7 2000		39015066073894	01	1



Open... Export ▼ Help

39015066073894 01

Facet / Filter Undo / Redo o

Infinite undo history

Don't worry about making mistakes. Every change you make will be shown here, and you can undo your changes anytime.

Learn more »

Extrac.

A S2. HATCH

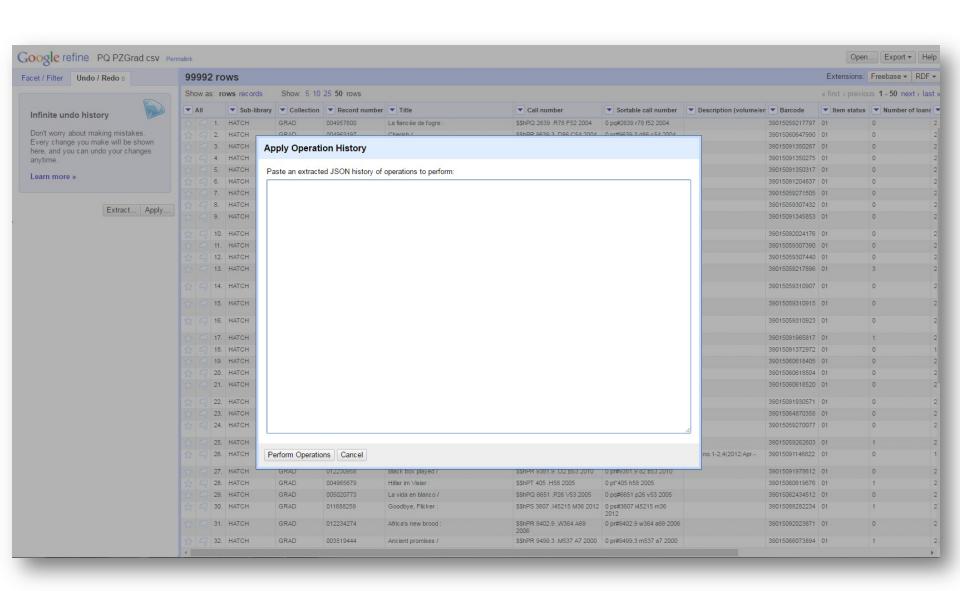
GRAD

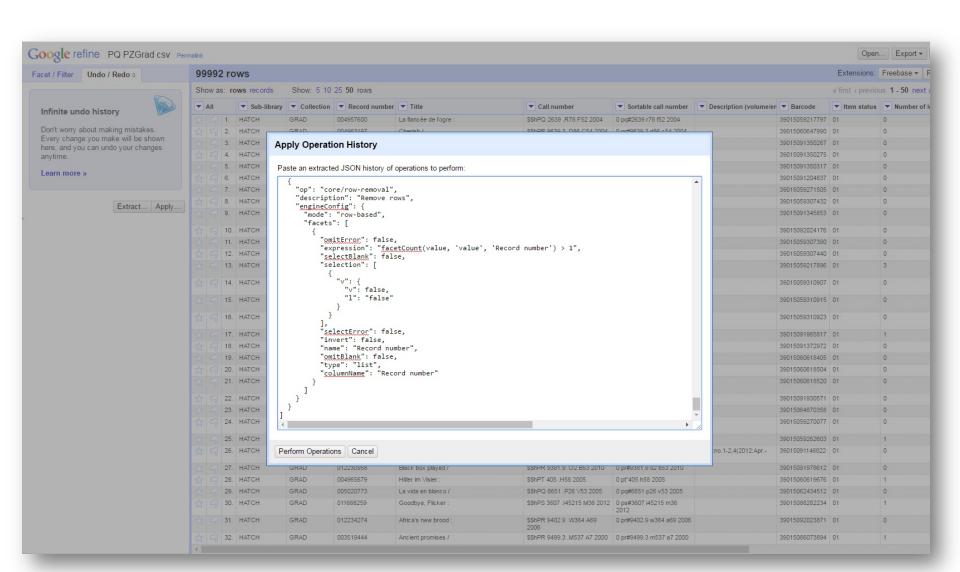
003519444

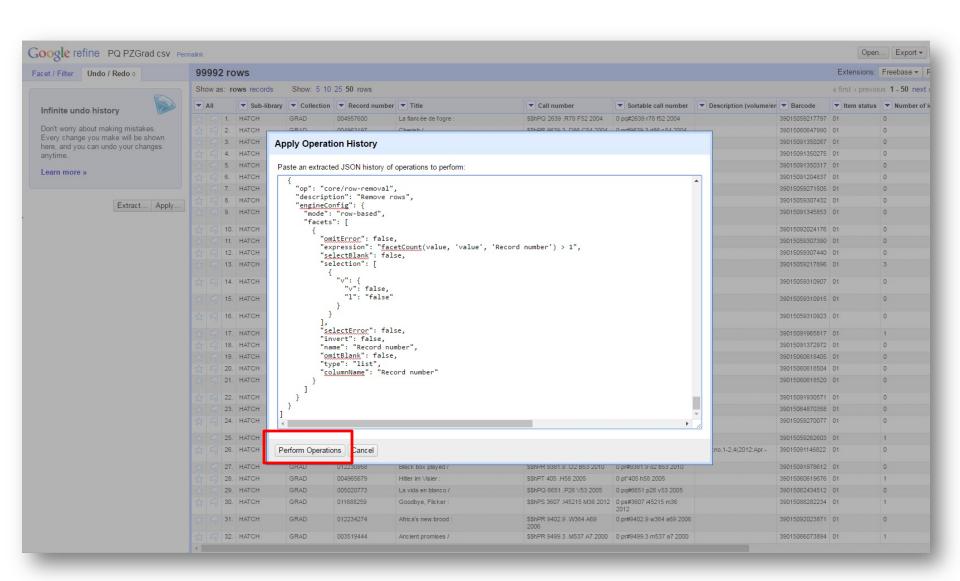
Ancient promises /

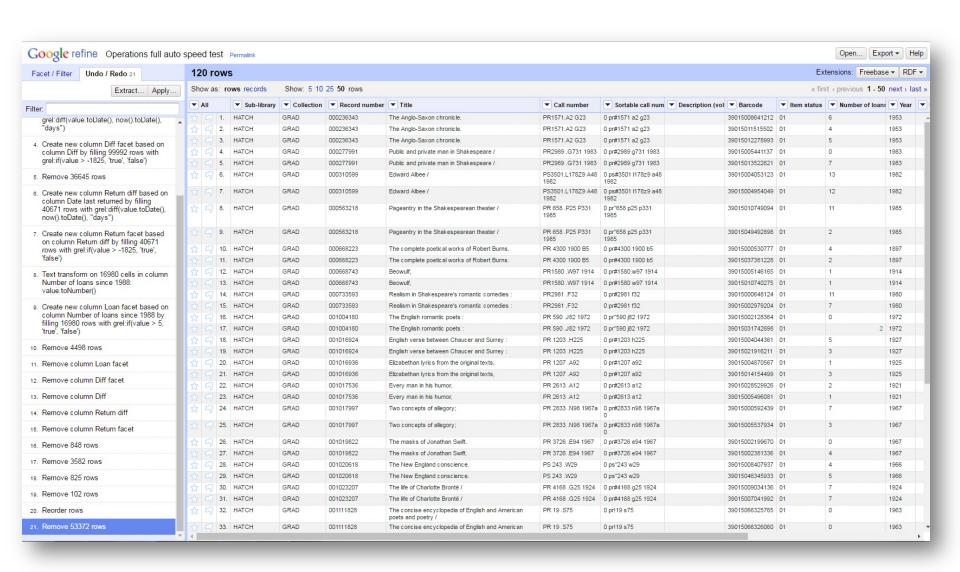
	999	92	ro	ws								Extensions:	Freebase ▼ RDF ▼	
	Show as: rows records			ows records	Show: 5 10 25 50 rows								« first « previous 1 - 50 next » last »	
	▼ A	II		▼ Sub-library	▼ Collection	Record number	Title	▼ Call number	Sortable call number	Description (volume/en	▼ Barcode	▼ Item status	▼ Number of loans	
			1.	HATCH	GRAD	004957600	La fiancée de l'ogre :	\$\$hPQ 2639 .R78 F52 2004	0 pq#2639 r78 f52 2004		39015059217797	01	0 2	
			2.	HATCH	GRAD	004963197	Cherish /	\$\$hPR 9639.3 .D86 C54 2004	0 pr#9639.3 d86 c54 2004		39015060647990	01	0 2	
			3.	HATCH	GRAD	011717137	Le dernier amour d'Arsène Lupin :	\$\$hPQ 2623 .E145 D47 2012	0 pq#2623 e145 d47 2012		39015091350267	01	0 2	
			4.	HATCH	GRAD	011717144	L'âme du monde /	\$\$hPQ 2672 .E5642 A67 2012	0 pq#2672 e5642 a67 2012		39015091350275	01	0 2	
			5.	HATCH	GRAD	011717152	Henri Meschonnic dans tous ses états :	\$\$hPQ 2673 .E712 Z73 2012	0 pq#2673 e712 z73 2012		39015091350317	01	0 2	
			6.	HATCH	GRAD	011660612	Melyador :	\$\$hPQ 1461 .F8 M34 2012	0 pq#1461 f8 m34 2012	v.2	39015091204837	01	0 2	
			7.	HATCH	GRAD	004374239	The sugar syndrome /	\$\$hPR 6116 .R42 S93 2003	0 pr#6116 r42 s93 2003		39015059271505	01	0 2	
- 7			8.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden:	\$\$hPT 2601 .N353 2004	0 pt#2601 n353 2004	v.6	39015059307432	01	0 2	
4			9.	HATCH	GRAD	011660767	Vivir sin los otros :	\$\$hPQ 8180.417 .0593 V58 2010	0 pq#8180.417 o593 v58 2010		39015091345853	01	0 2	
			10.	HATCH	GRAD	012234286	Voices in the wind :	\$\$hPR 9402.5 .V65 2010	0 pr#9402.5 v65 2010		39015092024176	01	0 2	
			11.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden:	\$\$hPT 2601 .N353 2004	0 pt#2601 n353 2004	v.3	39015059307390	01	0 2	
			12.	HATCH	GRAD	004919102	Gesammelte Werke in zehn Bänden:	\$\$hPT 2601 .N353 2004	0 pt#2601 n353 2004	v.7	39015059307440	01	0 2	
			13.	HATCH	GRAD	004957595	Albert Camus et les écritures algériennes :	\$\$hPQ 2605 .A3734 Z5 R44 2003	0 pq#2605 a3734 z5 r44 2003		39015059217896	01	3 2	
			14.	HATCH	GRAD	004942383	Los puentes de la traducción :	\$\$hPQ 7297 .P285 Z5 B73 2004	0 pq#7297 p285 z5 b73 2004		39015059310907	01	0 2	
			15.	HATCH	GRAD	004942384	Copyright /	\$\$hPQ 8098.26 .R25 C66 2003	0 pq#8098.26 r25 c66 2003		39015059310915	01	0 2	
			16.	HATCH	GRAD	004942386	Octavio Paz :	\$\$hPQ 7297 .P285 Z5 O28 2004	0 pq#7297 p285 z5 o28 2004		39015059310923	01	0 2	
			17.	HATCH	GRAD	012228383	Reflections & retrospectives /	\$\$hPR 9344 .R44 2012	0 pr#9344 r44 2012		39015091965817	01	1 2	
			18.	HATCH	GRAD	012229441	L'ombre de mémoire /	\$\$hPQ 2605 .O33 O43 1990	0 pq#2605 o33 o43 1990		39015091372972	01	0 1	
			19.	HATCH	GRAD	004955215	Cuore nella letteratura per l'infanzia /	\$\$hPQ 4683 .A3 C83 2004	0 pq#4683 a3 c83 2004		39015060618405	01	0 2	
			20.	HATCH	GRAD	004955217	Morso di luna nuova :	\$\$hPQ 4864 .E549 M67 2005	0 pq#4864 e549 m67 2005		39015060618504	01	0 2	
			21.	HATCH	GRAD	004955219	Da un'altra carne /	\$\$hPQ 4864 .E81228 D31 2004	0 pq#4864 e81228 d31 2004		39015060618520	01	0 2	
			22.	HATCH	GRAD	012231343	Pour l'histoire du "Thésée" d'André Gide :	\$\$hPQ 2613 .I2 T52 2012	0 pq#2613 i2 t52 2012		39015091930571	01	0 2	
			23.	HATCH	GRAD	005281827	Benchmarks of education /	\$\$hPR 9387.9 .K93 B46 2004	0 pr#9387.9 k93 b46 2004		39015064870358	01	0 2	
			24.	HATCH	GRAD	004966506	Domínios /	\$\$hPQ 9698.413 .A73 D66 2004	0 pq#9698.413 a73 d66 2004		39015059270077	01	0 2	
			25.	HATCH	GRAD	004955273	Falscher Hase :	\$\$hPT 2668 .E625 F35 2005	0 pt#2668 e625 f35 2005		39015059262603	01	1 2	
			26.	HATCH	GRAD	000677633	Poetry.	\$\$hPS 301 .P74	0 ps"301 p74	v.200:no.1-2,4(2012:Apr Aug.)	39015091146822	01	0 1	
			27.	HATCH	GRAD	012230958	Black box played /	\$\$hPR 9381.9 .O2 B53 2010	0 pr#9381.9 o2 b53 2010		39015091978612	01	0 2	
			28.	HATCH	GRAD	004965679	Hitler im ∀isier :	\$\$hPT 405 .H58 2005	0 pt"405 h58 2005		39015060619676	01	1 2	
			29.	HATCH	GRAD	005020773	La vida en blanco /	\$\$hPQ 6651 .P26 V53 2005	0 pq#6651 p26 v53 2005		39015062434512	01	0 2	
			30.	HATCH	GRAD	011688259	Goodbye, Flicker:	\$\$hPS 3607 .145215 M36 2012	0 ps#3607 i45215 m36 2012		39015088282234	01	1 2	
			31.	HATCH	GRAD	012234274	Africa's new brood :	\$\$hPR 9402.9 .W364 A69	0 pr#9402.9 w364 a69 2006		39015092023871	01	0 2	

\$\$hPR 9499.3 .M537 A7 2000 0 pr#9499.3 m537 a7 2000









Outcomes

- Cut processing time by approximately 99.8%.
- Cut weekly staff time on project from 34 hours to 7 hours.
- Processed over 2 million items.
- Withdrawn 9,000 low-use duplicates from our collections.

Questions?

OpenRefine command scripts on Github:

https://github.com/mcarruthers/Operations_dedup

mcarruth@umich.edu @mattadata2