Sign Up | Sign In

Columbia Journalism School 🕁

Tutorial Merging two tables in Open Refine

- The data:

Go to Major Contract Awards: <u>https://finances.worldbank.org/Procurement/Major-Contract-Awards/kdui-wcs3</u>

Filter by Region: Latin America

Go to Export and choose CSV to export all the data in csv format.

(R) WORLD BANK GROUP	IBRD	IDA	IFC	MIGA	ICSID	
N I I WORLED BRUIK ORGON						

World Bank Group Finances

Ma	jor Contract A	wards				2	f y 🖂 🗮	🗄 🗄 🗖 🔍 Find in this Dataset
This	s set of contract aw	vards includes data on o	commitments against contra	acts that were reviewed by the l	Bank 🕨 💽	Kanage 🐼 More View	🕫 🍸 Filter 🔛 V	isualize 🕞 Export 📴 Discuss 🚺 Embed 👩 About
	As of Date	Fiscal Year	Region	Borrower Country 10 🗎	Borrower Country Code	Project ID 🚯 🗄	Project Name	Export ×
1 :=	Oct 31, 2015	2000	Africa	Africa	34	P066627	Strengthen Canacit	SODA API
2 ==	Oct 31, 2015	2000	Africa	Africa	24	P000027		OData d
2 :=	001 31, 2015	2000	Africa	Africa	34	P000003	REINF (CEN.ENV.I	Print
3 :=	Oct 31, 2015	2000	Africa	Africa	3A	P000017	3A-1G/BN Enginee	Developed
4 :≡	Oct 31, 2015	2000	Africa	Africa	3A	P000001	West Africa Pilot Co	Download
5 🗮	Oct 31, 2015	2000	Africa	Africa	3A	P000001	West Africa Pilot Co	Download a copy of this dataset in a static
6 등	Oct 31, 2015	2000	Africa	Africa	3A	P000001	West Africa Pilot Co	format
7 등	Oct 31, 2015	2000	Africa	Africa	3A	P000001	West Africa Pilot	LOWINGLE A
8 🗮	Oct 31, 2015	2000	Africa	Africa	3A	P000001	West Africa P	CSV
9 ;≣	Oct 31, 2015	2000	Africa	Africa	3A	P000003	REIMP(CEN.ENV.	in the second seco
10 🗮	Oct 31, 2015	2000	Africa	Africa	3A	P052301	African Virtual Univ	ISON
11 🖂	Oct 31, 2015	2000	Africa	Africa	3A	P000003	REIMP(CEN.ENV.I	
12 🗮	Oct 31, 2015	2000	Africa	Africa	3A	P069721	COMESA: Creation	PDF
13 \Xi	Oct 31, 2015	2000	Africa	Africa	3A	P000003	REIMP(CEN.ENV.I	RDF
14 🗄	Oct 31, 2015	2000	Africa	Africa	3A	P000003	REIMP(CEN.ENV.I	RSS
15 🗮	Oct 31, 2015	2000	Africa	Africa	3A	P000003	REIMP(CEN.ENV.I	XLS
16 :=	Oct 31 2015	2000	Africa	Africa	34	P00003	REIMPICEN ENVI	XLSX

Columbia Journalism School 🛥

USE THE CLEAN FILE FROM http://ieg.worldbankgroup.org/data

2- Open the two files in separate browser tabs as shown below: Open OpenRefine from the browser by typing http://127.0.0.1:3333 Select file: Major_Contract_Awards.csv Click Next CLICK Create Project.

Refine Major_Contract_Awards csv Permalink Open Export + Hel												Export - Help	
Facet / Filter Undo / Redo 0	20	205340 rows Extensions: undefin											
	Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 10 next > last »												
Using facets and filters				As of Date	Fiscal Year	Fiscal Year 💌 Region 💌 Borrower Count 💌 Borrower Count 💌 Pro	💌 Project ID	Project Name	Procurement Ty				
Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top			1.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	P066627	Strengthen Capacity of ACBF to Manage & Operate the PACT	Implementation Activity	Consultant Services	
of each data column.	☆		2.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	P000003	REIMP(CEN.ENV.INFO)	Implementation Activity	Consultant Services	
Watch these screencasts			3.	10/31/2015 12:00:00 AM	2000	Africa	Africa	ЗА	P000017	3A-TG/BN Engineering TAL (FY92)	Equipment Information Technology	Goods	
	☆		4.	10/31/2015 12:00:00 AM	2000	Africa	Africa	ЗA	P000001	West Africa Pilot Community-based Natural Resources & Wildli	Feasibility studies, Design Studies	Consultant Services	
	☆		5.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	P000001	West Africa Pilot Community-based Natural Resources &	Feasibility studies, Design Studies	Consultant Services	

Open another tab

(type http://127.0.0.1:3333 in the new tab to open OpenRefine)

Open file: IEG_World_Bank_Project_Performance_Ratings_asof_Jan182017_download xlsx

3- Go to the file "Major_Contract_Awards.csv" and find column Project ID Click Edit Columns "Add Column based on this column"

Columbia Journalism School ₪	Merging two tables in Open Refine									gs	Gi 27990	iannina (@colum	Seg bia.	nini <u>edu</u>
Refine Major_Contract_Aw	ards	csv	Pe	rmalink								Open	-AP-	Help
Facet / Filter Undo / Redo 0	20)53	40 ı	rows								Extension	s:efi	ned 🕶
	Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 10 next > last »													
Lising facets and filters		All		As of Date	Fiscal Year	Region	Borrower Count	a 💌 Borrower Count	Project ID	💌 Pro	oject Name	Procurement Type	Proc	urement C
Use facets and filters to select subsets	☆		1.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	Facet Text filter	•	en Capacity of Manage & the PACT	Implementation Activity	Consultar	t Services
filter methods from the menus at the top									Edit cells	•	-			
of each data column.	슔		2.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	Edit column	•	Split into seve	eral columns		Services
Not sure how to get started?				12.00.00741					Transpose	•	Add column b	ased on this column		
Watch these screencasts	53		3.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	Sort		Add column b	v fetching URLs		
									View	•		, ,		
	57		4.	10/31/2015	2000	Africa	Africa	3A	Reconcile	•	Rename this	column		Services
	\sim			12:00:00 AM						Natural	Remove this	column		
			_							Wildli	Move column	to beginning		
	হ্য		5.	10/31/2015 12:00:00 AM	2000	Africa	Africa	3A	P000001	West A Commu Natural Wildli	Move column Move column	to end left		Services
	슔		6.	10/31/2015	2000	Africa	Africa	3A	P000001	West A	f Move column	right		Services
				12:00:00 AM						Natural Wildli	Resources &	Design Studies		_

a- Name the new column as OUTCOME

b- Go to the empty box and replace the word "value" with the following script

if (value! ='null',cell.cross("IEG_World_Bank _Project_Performance_Ratings_asof_J an182017_download xlsx", "Project ID").cells["IEG_Outcome"].value [0],'')

Columbia Journalism School 🛥