



Handleiding OpenRefine

Deel 1 — Het begin

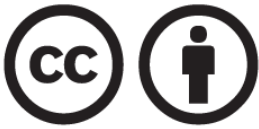
Nynke Kuipers—Kubus erfgoed & digitalisering

Oktober 2024



KUBUS.fri
erfgoed & digitalisering

Beeld voorpagina: Portret van mevrouw, Marie Jeannette de Lange door Jan Toorop, Publiek domein



Licentie

Op deze handleiding is een Creative Commons Naamsvermelding-licentie van toepassing. Deze publicatie mag worden bewerkt, doorgestuurd, geüpload, gedownload, gekopieerd en anderszins verspreid, zolang de naam van de maker vermeld wordt. De gehele licentietekst is te vinden op:

<https://creativecommons.org/licenses/by/4.0/>

Als je deze handleiding wil hergebruiken kun je de volgende methode van naamsvermelding aanhouden:

Handleiding OpenRefine - deel 1 - het begin (Nynke Kuipers, Kubus erfgoed & digitalisering) CC-BY



Deze handleiding is tot stand gekomen binnen de Datawerkplaats van Landschap Erfgoed Utrecht.



Inhoudsopgave

Introductie	4
OpenRefine installeren	5
Download	5
Online versie	5
Project maken in OpenRefine	6
Een bestaand project openen in OpenRefine	7
Kolommen verplaatsen in OpenRefine	8
Optie 1: één kolom verplaatsen.....	8
Optie 2: volgorde van alle kolommen veranderen.....	9
Rijen en records	10
Cellen splitsen en samenvoegen	11
Cellen splitsen	11
Cellen samenvoegen.....	14
Kolommen samenvoegen en splitsen	16
Kolommen samenvoegen	16
Kolommen splitsen.....	18
Undo / redo	20
Zoeken in OpenRefine	21
Facet zoeken	21
Tekst zoeken.....	22
Facet resultaten opslaan.....	23
Facetten combineren.....	24
Facet op kolommen.....	26
Links en verwijzingen	27

Introductie

OpenRefine is een open source tool specifiek ontwikkeld om met ‘messy data’, ofwel: rommelige data, te werken. Het is ook een prachtige tool om op een andere manier naar je erfgoeddata te kijken. Deze handleiding neemt je mee door een aantal basis functies in OpenRefine waarmee je met je eigen erfgoeddata aan de slag kan. Het leert je hoe je OpenRefine kunt inzetten in het **bekijken** en **beoordelen** van je eigen data.

De handleiding is opgebouwd aan de hand van een oefenbestand, waarmee enkele functies van OpenRefine getoond worden. Het oefenbestand is een willekeurig samengestelde csv, specifiek samengesteld is om de mogelijkheden van OpenRefine te kunnen tonen. Alle data is gefingeerd en kunnen geen rechten aan ontleend worden. Om ook zelf de stappen uit de handleiding te kunnen reproduceren is het bestand te downloaden: [oefenbestand \(csv\)](#)

In deze handleiding is gewerkt met [OpenRefine Version 3.8.2](#). Bij nieuwere of oudere versies kan het er net iets anders uitzien dan getoond in de screenshots van deze handleiding.

Deze handleiding is ontwikkeld naar aanleiding van workshops die gegeven zijn in het kader van de Datawerkplaats Utrecht, een initiatief van Landschap Erfgoed Utrecht en Netwerk Digitaal Erfgoed.

Voor uitgebreide instructies over het gebruik van OpenRefine is er de [user manual](#) (engelstalig) van OpenRefine zelf.

OpenRefine installeren

Voordat je aan de slag kunt, dien je OpenRefine te downloaden: <https://openrefine.org/>. Je installeert OpenRefine lokaal op de computer. Je hebt geen internet nodig om met OpenRefine te werken, al lijkt het anders omdat het in een internetbrowser opent. Je hebt alleen internet nodig als je bijvoorbeeld met reconciliation met online bronnen bezig gaat. Je projecten worden lokaal op de computer opgeslagen en niet in een cloud. Hierdoor kan je ook goed met gevoelige data aan de slag. Overleg met je ICT-dienstverlener als je niet zelf programma's mag installeren.

Download

OpenRefine is gratis te downloaden. [Hier](#) vind je de laatste versie van het programma en de verschillende downloads voor Windows, Mac OS en Linux. Let op: Om OpenRefine in Windows te kunnen gebruiken dien je ook Java Runtime Environment te installeren. Als dit nog niet is gebeurd, kies dan het downloadbestand inclusief Java.

- Wanneer de download is voltooid kun je het ZIP-bestand uitpakken.
- Vervolgens open je de Toepassing: openrefine
- Het programma zal nu openen in de voorkeurs-browser.

Online versie

Niet gelukt om OpenRefine aan de praat te krijgen op je computer? Maar wil je toch snel van start? Dankzij het Netwerk Digitaal Erfgoed is er een online toegang tot een OpenRefine applicatie:

<https://openrefine.netwerkdigitaalerfgoed.nl/>

Let op: de projecten zijn voor alle gebruikers van deze link te bekijken, werk hier niet met gevoelige data

Let op: de opgeslagen projecten worden periodiek verwijderd.

Project maken in OpenRefine

OpenRefine kan met diverse export-formaten / bestandsformaten uit de voeten. TSV, CSV, Excel, XML, RDF, etc.

1. Ga naar **bladeren** om het gewenste bestand te selecteren > **next**

OpenRefine A power tool for working with messy data.

Create project

Create a project by importing data. What kinds of data files can I import?
TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents are all supported. Support for other formats can be added with OpenRefine extensions.

Get data from

This Computer

Web Addresses (URLs)

Clipboard

Database

Google Data

Locate one or more files on your computer to upload:

Bladeren... Oefenbestand-OpenRefine - het begin.csv

Next »

2. Instellingen

OpenRefine ziet zelf al wat voor soort bestand het is en maakt een **preview**. Je kunt de instellingen zodanig aanpassen totdat je preview het gewenste resultaat heeft.

(Voor de beginner: niets veranderen werkt vaak al prima).

OpenRefine A power tool for working with messy data.

Create project

« start over Configure parsing options Project name: Oefenbestand OpenRefine het beg Tags

Create project »

	Vervaardiger	Objectnummer	Soort object	Materiaal	Datering	Afmeting type	Afmeting waarde	Afmeting eenheid	Titel	Beschrijving	Auteursrec
1.	Rembrandt	ON0003.50001	schilderij	doek/olieverf	19-jun-04	hoogte breedte diepte	10.00 225.00 525.00	mm mm mm			CC0
2.	Rembrandt	ON0003.50002	ets		1632						CC0
3.	Rembrandt	ON0003.50003	ets		1831	breedte	380.00	mm			CC0
4.	Rembrandt	ON0003.50004	ets		1767	hoogte breedte	1,540.00 620.00	mm mm			CC0
5.	Rembrandt	ON0003.50005	schilderij	olieverf paneel	11/8/1870	hoogte breedte diepte	15.40 7.50 8.80	cm cm cm	Rembrandt	Portret van een man met een baret, borststuk en face. Naar een zelfportret van Rembrandt. In bruine lijst met goudkleurige binnenlijst.	CC0
6.	Rembrandt	ON0003.50006	ets								CC0

Parse data as

Character encoding: WINDOWS-1252

Update preview

Disable auto preview

CSV / TSV / separator-based files

Line-based text files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

JSON-LD files

RDF/JSON files

RDF/JSON-Triples files

Columns are separated by

commas (CSV)

tabs (TSV)

custom ;

Use character " to enclose cells containing column separators

Trim leading & trailing whitespace from strings

Escape special characters with \

Ignore first 0 line(s) at beginning of file

Parse next 1 line(s) as column headers

Column names (comma separated)

Discard initial 0 row(s) of data

Load at most 0 row(s) of data

Attempt to parse cell text into numbers

Store blank rows

Store blank cells as nulls

Store file source

Store archive file

3. Create project

Een bestaand project openen in OpenRefine

Als je lokaal werkt, worden je projecten op je computer opgeslagen. Je kunt zo altijd weer verder werken in een bestaand project.

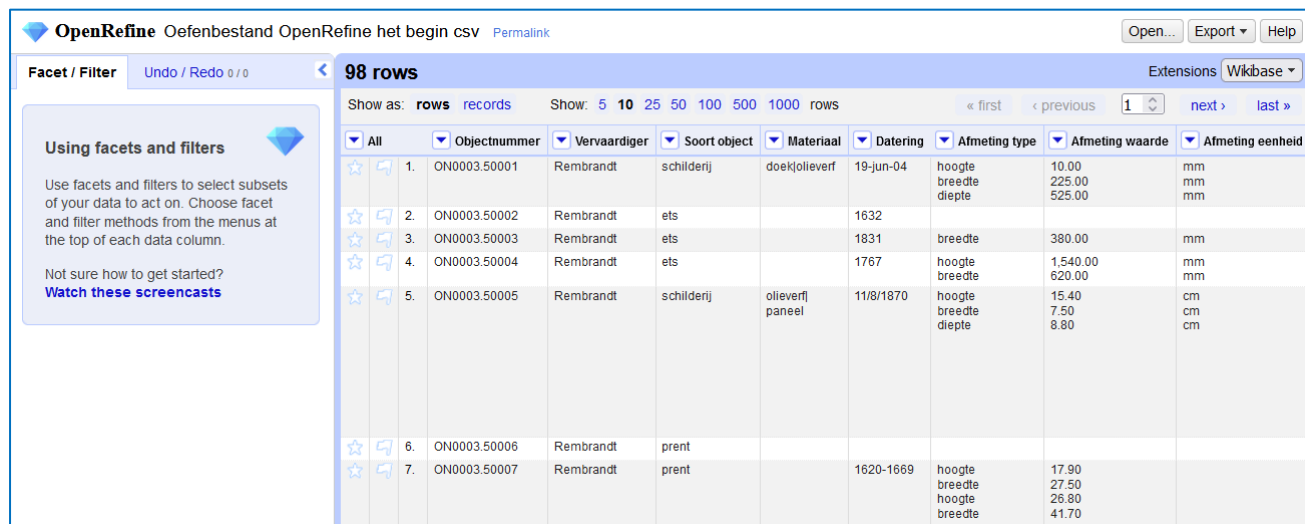
1. Ga naar **Open project** > klik op het project dat je wilt openen



The screenshot shows the OpenRefine application window. On the left, a sidebar menu is visible with the following options: 'Create project', 'Open project' (highlighted with a red box and a red number 1), 'Import project', and 'Language settings'. The main area displays a table of existing projects. The table has columns for 'Last modified', 'Name', 'Tags', 'Creator', 'Subject', 'Description', and 'Row Count'. The projects listed are:

Last modified	Name	Tags	Creator	Subject	Description	Row Count
Oct 26, 2024, 1:11:43 PM	OpenRefine de basis Axiell demo afmetingen csv					121
Oct 26, 2024, 1:11:33 PM	OpenRefine de basis Adlib inhoud titel en beschrijving csv					34
Oct 26, 2024, 1:11:21 PM	Oefenbestand OpenRefine het begin csv	cursus				98
Oct 26, 2024, 12:44:44 PM	OpenRefine de basis Axiell demo afmetingen csv	cursus				121

2. Je bent nu weer waar je gebleven was



The screenshot shows the OpenRefine application window with the project 'Oefenbestand OpenRefine het begin csv' open. The interface includes a 'Facet / Filter' sidebar on the left, a 'Using facets and filters' section, and a main data table. The table has 98 rows and columns for 'Objectnummer', 'Vervaardiger', 'Soort object', 'Materiaal', 'Datering', 'Afmeting type', 'Afmeting waarde', and 'Afmeting eenheid'. The data is displayed in a table format with various filters and sorting options available.

Objectnummer	Vervaardiger	Soort object	Materiaal	Datering	Afmeting type	Afmeting waarde	Afmeting eenheid
1. ON0003.50001	Rembrandt	schilderij	doek/olieverf	19-jun-04	hoogte breedte diepte	10.00 225.00 525.00	mm mm mm
2. ON0003.50002	Rembrandt	ets		1632			
3. ON0003.50003	Rembrandt	ets		1831	breedte	380.00	mm
4. ON0003.50004	Rembrandt	ets		1767	hoogte breedte	1,540.00 620.00	mm mm
5. ON0003.50005	Rembrandt	schilderij	olieverf paneel	11/8/1870	hoogte breedte diepte	15.40 7.50 8.80	cm cm cm
6. ON0003.50006	Rembrandt	prent					
7. ON0003.50007	Rembrandt	prent		1620-1669	hoogte breedte hoogte breedte	17.90 27.50 26.80 41.70	

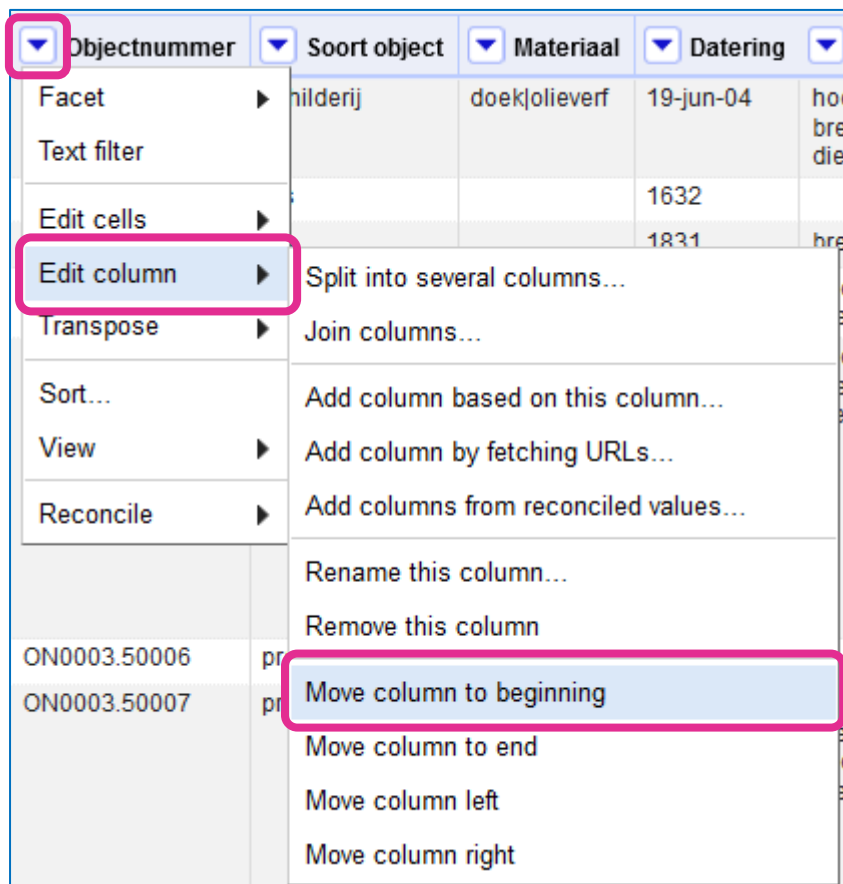
Kolommen verplaatsen in OpenRefine

Optie 1: één kolom verplaatsen

Om optimaal de functionaliteiten van OpenRefine te gebruiken, zorg altijd voor een unieke *record identifier* als eerste kolom (bijvoorbeeld objectnummer of recordnummer).

In dit voorbeeld wil ik de kolom “objectnummer” helemaal naar het begin verplaatsen

1. Klik op het pijltje voor de kolomnaam > **Edit column** > **Move column to beginning**

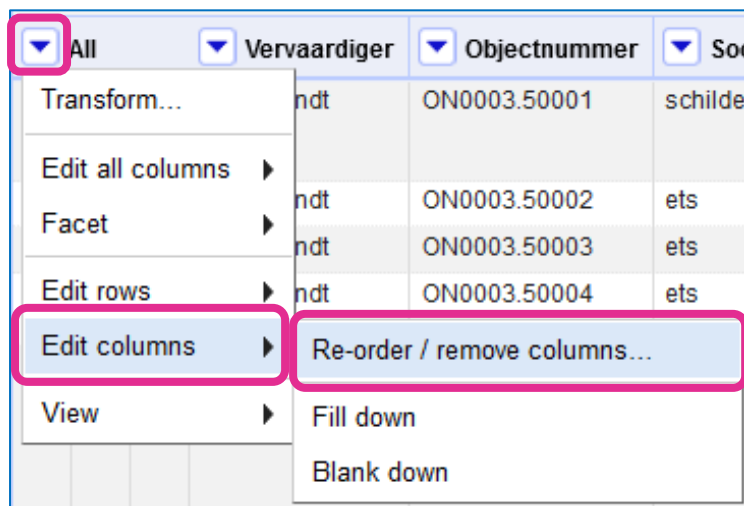


Ook leuk: Je ziet dat er meer opties zijn, zo kan je een kolom ook één plaatsje naar links of rechts verplaatsen, of juist helemaal naar het eind.

Optie 2: volgorde van alle kolommen veranderen

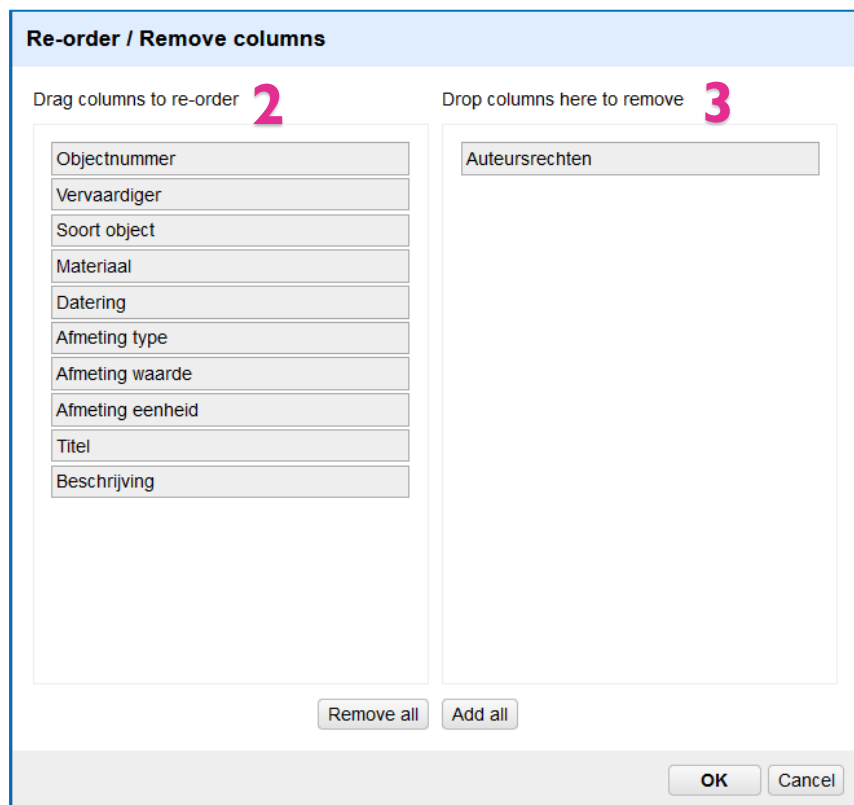
Zijn er meerdere kolommen waarvan je de volgorde wilt aanpassen?

1. Klik op het pijltje voor de kolom **All** > **Edit columns** > **Re-order / remove columns...**



2. Je kunt nu de blokjes met de kolomnamen naar de gewenste plek schuiven (drag and drop).

3. Als je kolommen wilt verwijderen drop je ze in de rechter helft.



Rijen en records

OpenRefine maakt onderscheid in de weergave in rijen en records. Dit betekent dat je in een recordweergave velden met meerdere regels (occurrences) kan tonen en die blijven dan ook bij het record. Bekijk het volgende voorbeeld met afmetingen. Let op het **aantal** en de **‘rijnummers’**.

Records

98 records

Show as: rows records

Show: 5 10 25 50 100 500 1000 records

« first < previous 1

All	Objectnummer	Vervaardiger	Soort object	Materiaal	Datering	Afmeting type	Afmeting waarde	
☆	1.	ON0003.50001	Rembrandt	schilderij	doek olieverf	19-jun-04	hoogte	10.00
☆						breedte	225.00	
☆						diepte	525.00	
☆	2.	ON0003.50002	Rembrandt	ets		1632		
☆	3.	ON0003.50003	Rembrandt	ets		1831	breedte	380.00
☆	4.	ON0003.50004	Rembrandt	ets		1767	hoogte	1,540.00
☆						breedte	620.00	
☆	5.	ON0003.50005	Rembrandt	schilderij	olieverf paneel	11/8/1870	hoogte	15.40
☆						breedte	7.50	
☆						diepte	8.80	
☆	6.	ON0003.50006	Rembrandt	prent				

Rijen

149 rows

Ext...

Show as: rows records

Show: 5 10 25 50 100 500 1000 rows

« first < previous 1

All	Objectnummer	Vervaardiger	Soort object	Materiaal	Datering	Afmeting type	Afmeting waarde	
☆	1.	ON0003.50001	Rembrandt	schilderij	doek olieverf	19-jun-04	hoogte	10.00
☆	2.						breedte	225.00
☆	3.						diepte	525.00
☆	4.	ON0003.50002	Rembrandt	ets		1632		
☆	5.	ON0003.50003	Rembrandt	ets		1831	breedte	380.00
☆	6.	ON0003.50004	Rembrandt	ets		1767	hoogte	1,540.00
☆	7.						breedte	620.00
☆	8.	ON0003.50005	Rembrandt	schilderij	olieverf paneel	11/8/1870	hoogte	15.40

Cellen splitsen en samenvoegen

We moeten nu nog een stap zetten voordat we echt aan de slag kunnen. In het oefenbestand staan de afmetingen weergegeven als herhaalbare velden, maar deze zijn in 1 cel geplaatst. Dit willen we eerst graag splitsen zodat de informatie verdeeld is over meerdere cellen. (Deze weergave van csv is overeenkomstig met een csv-export uit Axiell Collections).

Cellen splitsen

We gaan eerst zorgen dat de data niet meer 'onder elkaar' staat in 1 cel.

1. Klik op het pijltje voor de kolomnaam waar je de actie op wilt uitvoeren. In dit voorbeeld Afmeting type > **Edit cells** > **Common transforms** > **Collapse consecutive whitespace**

Afmeting type	Afmeting waarde	Afmeting eenheid	Titel	Beschrijving	Au
Facet	00	mm			CC0
Text filter	.00	mm			
	.00	mm			
Edit cells	Transform...				CC0
Edit column	Common transforms	Trim leading and trailing whitespace			CC0
Transpose	Fill down	Collapse consecutive whitespace			
Sort...	Blank down	Unescape HTML entities			
View	Split multi-valued cells...	Replace smart quotes with ASCII			
Reconcile	Join multi-valued cells...	To titlecase			
	Cluster and edit...	To uppercase			
	Replace...	To lowercase			
hoogte breedte	17.90	To number			
hoogte breedte	27.50	To date			
	26.80	To text			
	41.70	To null			
		To empty string			

2. De data is nu achter elkaar geplaatst in plaats van onder elkaar

☒ **Afmeting type**

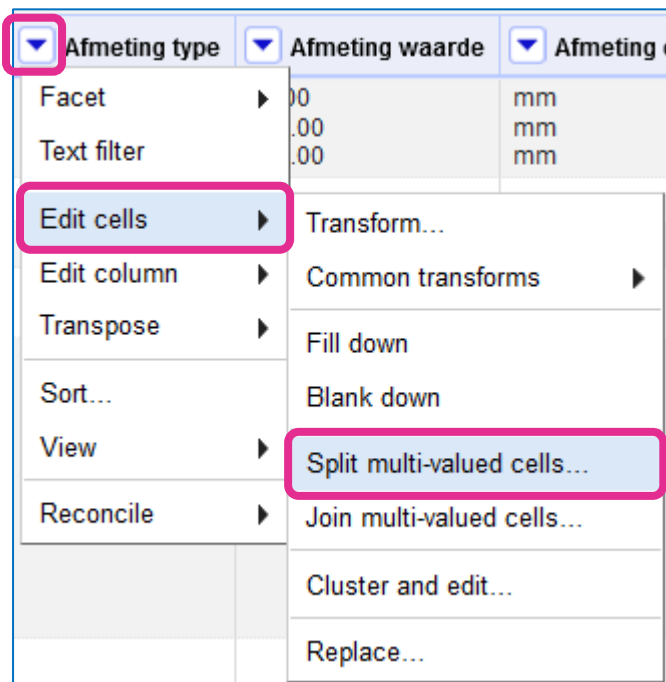
hoogte breedte
diepte

breedte

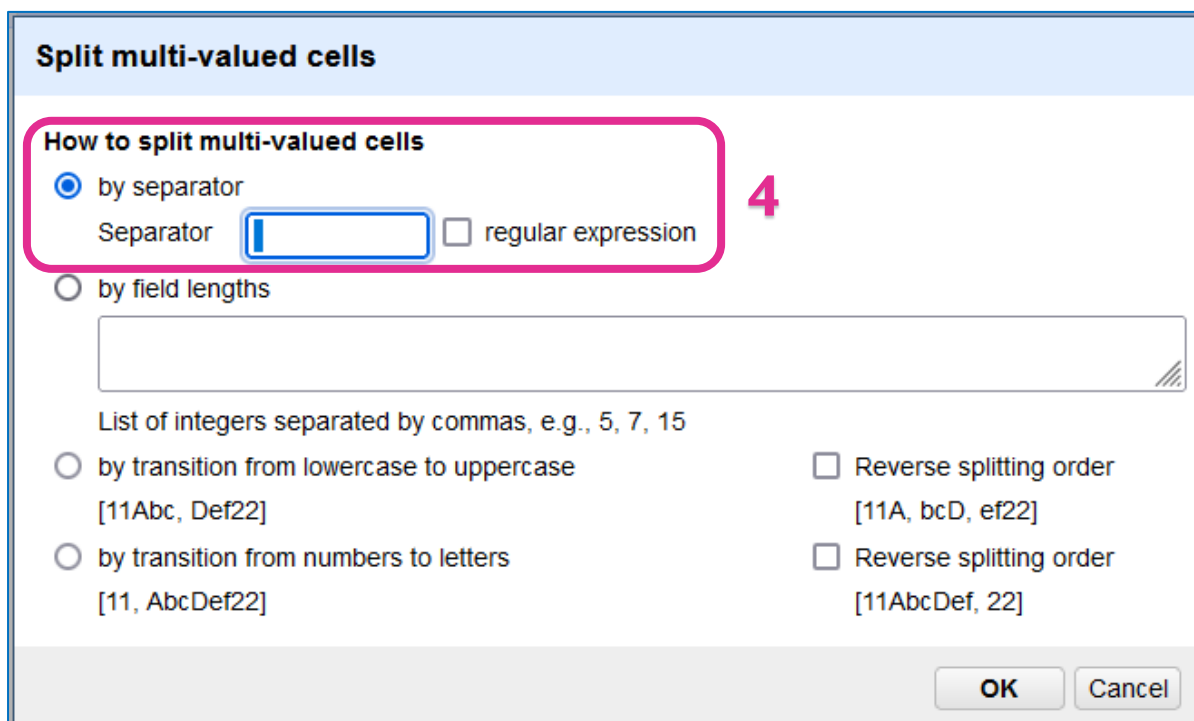
hoogte breedte

hoogte breedte
diepte

3. Klik wederom op het pijltje voor de kolomnaam > **Edit cells** > **Split multi-valued cells...**



4. Voer nu het scheidingsteken in, de **separator**. In dit geval is het een **spatie** > druk op **OK**



5. Het resultaat: de data is nu gesplitst over meerdere cellen (rijen)

▼ Afmeting type	▼ Afmeting waarde	▼ Afmeting eenheid
hoogte	10.00	mm
	225.00	mm
	525.00	mm
breedte		
diepte		
breedte	380.00	mm
hoogte	1,540.00	mm
	620.00	mm
breedte		
hoogte	15.40	cm
	7.50	cm
	8.80	cm
breedte		
diepte		

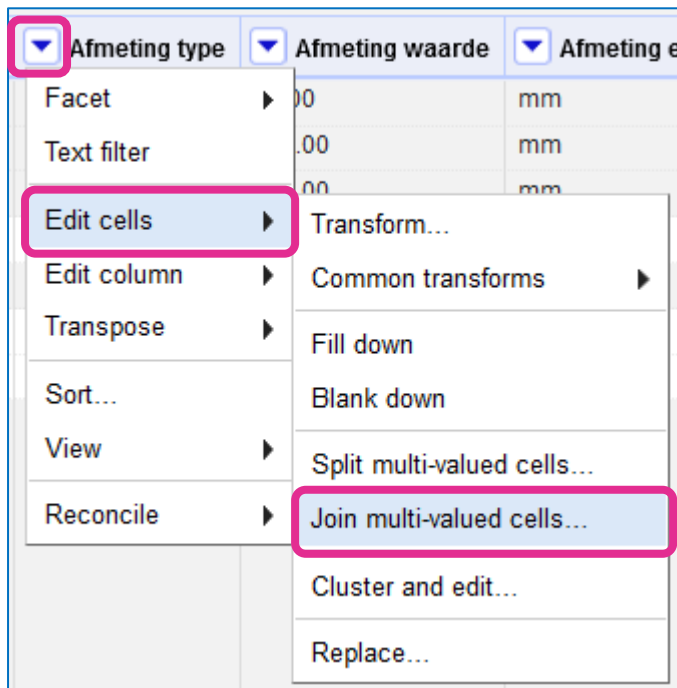
6. Voer nu de stappen 1 t/m 5 opnieuw uit voor de kolommen **Afmeting waarde** en **Afmeting eenheid**

▼ Afmeting type	▼ Afmeting waarde	▼ Afmeting eenheid
hoogte	10.00	mm
breedte	225.00	mm
diepte	525.00	mm
breedte	380.00	mm
hoogte	1,540.00	mm
breedte	620.00	mm
hoogte	15.40	cm
breedte	7.50	cm
diepte	8.80	cm

Cellen samenvoegen

Net zoals we cellen hebben gesplitst, kunnen we deze ook weer samenvoegen tot 1 cel.

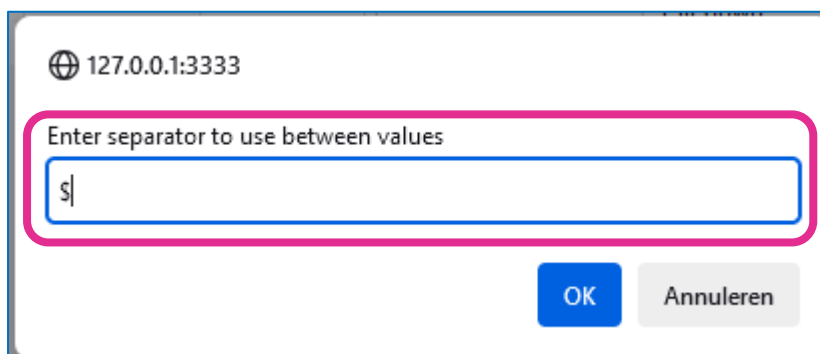
1. Klik op het pijltje voor de kolom waar je deze actie op wilt uitvoeren > **Edit cells** > **Join multi-valued cells**



2. Kies het scheidingsteken – de **seperator**

Kies hiervoor bij voorkeur een weinig gebruikt teken in je data. Als je een teken gebruikt dat al in je data voorkomt, ontstaan er fouten bij het weer splitsen van de data.

Gebruik bijvoorbeeld dollartekenen \$ of staand streepje | > klik op **OK**



3. Je data uit meerdere velden is nu samengevoegd met het dollarteken in één cel

▼ Afmeting type	▼ Afmeting waarde	▼ Afmeting eenheid
hoogte\$breedte\$diepte	10.00	mm
	225.00	mm
	525.00	mm
breedte	380.00	mm
hoogte\$breedte	1,540.00	mm
	620.00	mm
hoogte\$breedte\$diepte	15.40	cm
	7.50	cm
	8.80	cm

Kolommen samenvoegen en splitsen

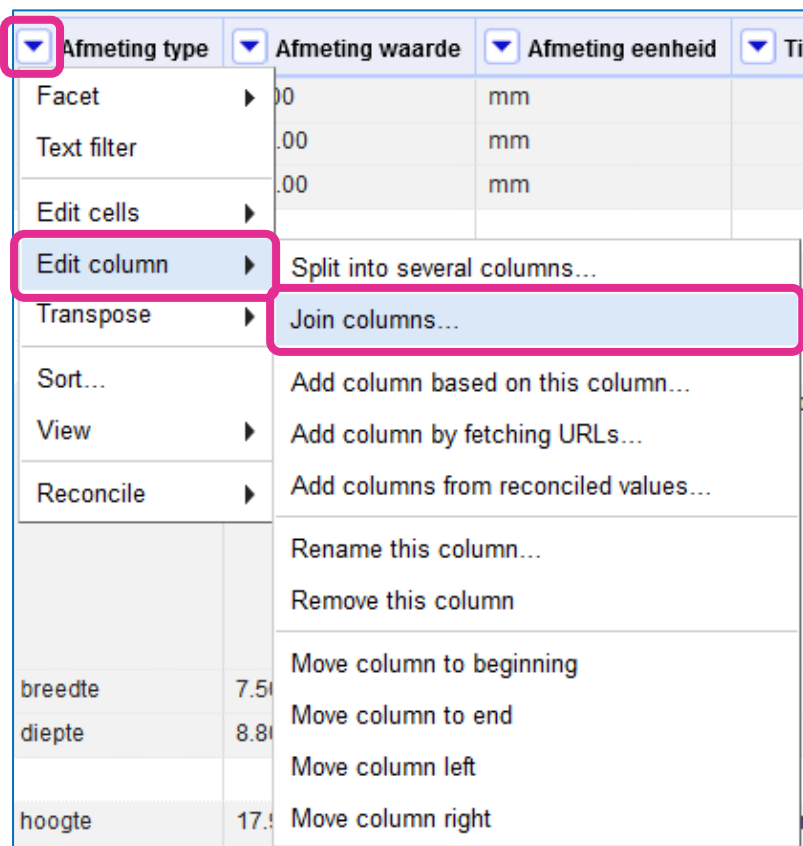
Ook kolommen kan je samenvoegen of splitsen, naar gelang wat je wilt bereiken. Dit kan een fijne actie zijn als er verschillende informatie zit in verschillende velden, die je wilt samenvoegen naar één veld.

Je kunt uiteindelijk ook de stappen combineren van kolommen en rijen: bijvoorbeeld als je meerdere occurrences wilt uitspreiden over meerdere kolommen in plaats van meerdere rijen: **Join multi valued cells > Split into several columns**.

Zo zijn er legio combinaties die je kunnen helpen bij het her-orderen van je data.

Kolommen samenvoegen

1. Klik op het pijltje voor de eerste kolom die je wil samenvoegen > **Edit Column > Join columns**



2. Selecteer de kolommen die je wilt samenvoegen

Join columns

Select and order columns to join **2**

- ☒ Afmeting type
- ☐ Objectnummer
- ☐ Vervaardiger
- ☐ Soort object
- ☐ Materiaal
- ☐ Datering
- ☒ Afmeting waarde
- ☒ Afmeting eenheid
- ☐ Titel

Select all Deselect all

Select options **3**

Separator between the content of each column \$ **a**
Enter one or more characters, or keep blank to join the columns without separator.

☐ Replace nulls with **b**
Enter one or more characters, or keep blank to replace nulls with blank strings.

☒ Skip nulls.

☐ In separator and nulls substitutes, use \n for new lines, \t for tabulation, \n for \n, \\t for \\t.

☐ Write result in selected column.

☒ Write result in new column named **c**

☐ Delete joined columns. **d**

OK Cancel

3. Loop vervolgens de instellingen na:

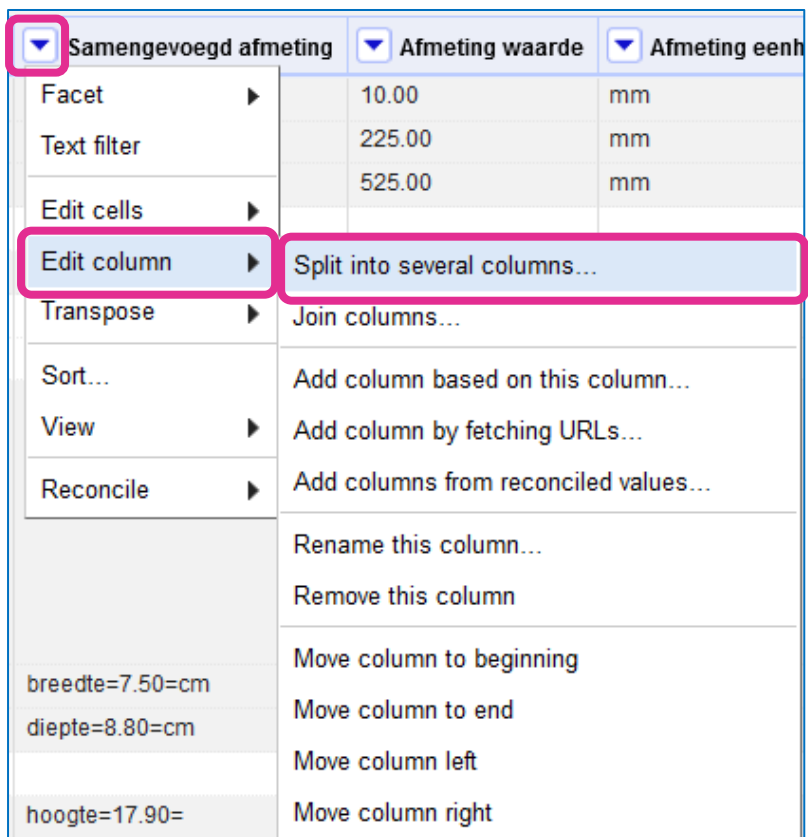
- Kies een scheidingsteken, de **separator**, in dit geval is een spatie het mooist, maar voor het voorbeeld is gekozen voor een \$ (dollar-teken)
- Wil je iets met de lege cellen in je samengevoegde kolom? In dit voorbeeld worden lege cellen overgeslagen
- Het resultaat wordt getoond in een nieuwe kolom, geef deze kolom een naam
- De oude kolommen wil ik in dit voorbeeld behouden, dus niet verwijderen

4. Het resultaat is een nieuw gevormde kolom met de volledige afmeting informatie bij elkaar gevoegd.

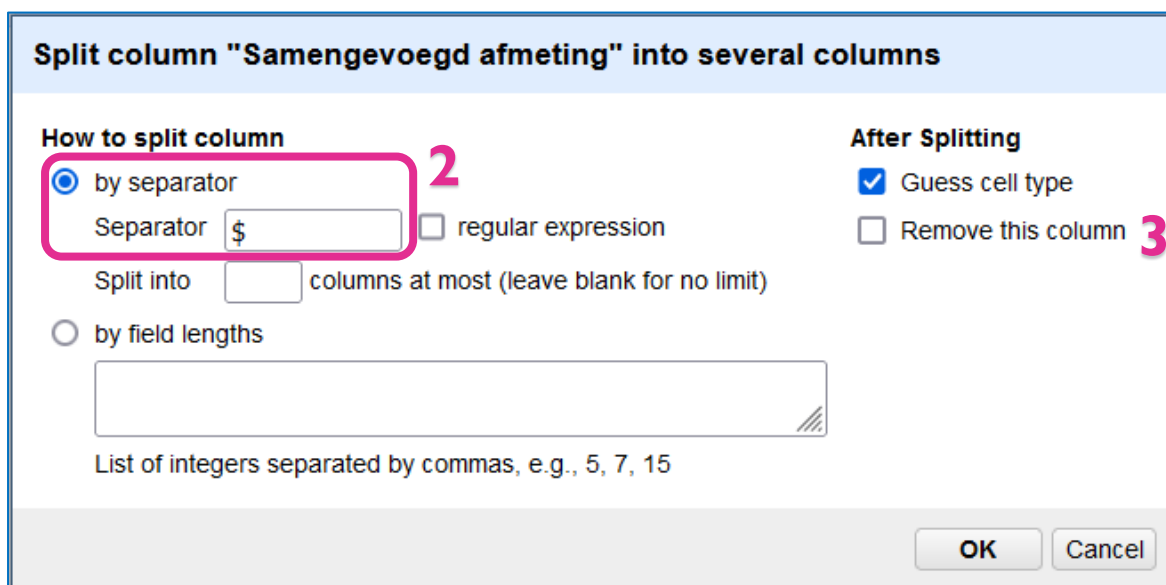
Afmeting type	Samengevoegd afmeting	Afmeting waarde	Afmeting eenheid
hoogte	hoogte\$10.00\$mm	10.00	mm
breedte	breedte\$225.00\$mm	225.00	mm
diepte	diepte\$525.00\$mm	525.00	mm
breedte	breedte\$380.00\$mm	380.00	mm
hoogte	hoogte\$1,540.00\$mm	1,540.00	mm
breedte	breedte\$620.00\$mm	620.00	mm
hoogte	hoogte\$15.40\$cm	15.40	cm

Kolommen splitsen

1. Klik op het pijltje voor de kolomnaam die je wilt splitsen > **Edit column** > **Split into several columns**



2. Voer de **seperator** in, in dit voorbeeld \$
3. Ook nu wil ik de oude kolom graag behouden



4. Er zijn nu drie nieuwe kolommen aangemaakt

▼ Samengevoegd afmeting	▼ Samengevoegd afmeting 1	▼ Samengevoegd afmeting 2	▼ Samengevoegd afmeting 3
hoogte\$10.00\$mm	hoogte	10	mm
breedte\$225.00\$mm	breedte	225	mm
diepte\$525.00\$mm	diepte	525	mm
breedte\$380.00\$mm	breedte	380	mm
hoogte\$1,540.00\$mm	hoogte	1,540.00	mm
breedte\$620.00\$mm	breedte	620	mm
hoogte\$15.40\$cm	hoogte	15.4	cm

5. OpenRefine heeft nu voor de kolom 'Samengevoegd afmeting 2' automatisch een 'getallen' kolom gemaakt. Als je dit niet wilt, kun je bij de instellingen bij het splitsen het vinkje uitzetten **Guess cell type**.

Undo / redo

Je kunt je aanpassingen in de tabel ongedaan maken. Zo kan je altijd iets uitproberen en simpelweg een stapje teruggaan om de actie ongedaan te maken.

Ook kan je al je stappen nalopen die je hebt gezet om je data tot een overzichtelijk en werkbaar tabel te maken. Handig voor als je dit vaker wilt doen en je niet meer weet hoe je daar ook alweer bent gekomen.

Facetten en zoekopdrachten zijn geen aanpassingen, dus die komen niet in je undo/redo lijst voor.

1. Ga naar het tabblad **Undo/Redo**
2. Klik op de stap waar je naar terug wilt > de stappen die ongedaan worden gemaakt worden gemarkeerd. In dit voorbeeld worden de laatste twee stappen teruggedraaid.

The screenshot shows the 'Undo / Redo' panel with 9 steps. The last step is highlighted in blue. The steps are:

0. Create project
1. Move column Objectnummer to position 0
2. Text transform on 35 cells in column Afmeting type: value.replace(/[\p{Zs}\s]+/, ' ')
3. Split multi-valued cells in column Afmeting type
4. Text transform on 35 cells in column Afmeting waarde: value.replace(/[\p{Zs}\s]+/, ' ')
5. Split multi-valued cells in column Afmeting waarde
6. Text transform on 35 cells in column Afmeting eenheid: value.replace(/[\p{Zs}\s]+/, ' ')
7. Split multi-valued cells in column Afmeting eenheid
8. Create new column Samengevoegd afmeting based on column Afmeting type by filling 149 rows with join ([cells['Afmeting type'].value,cells['Afmeting waarde'].value,cells['Afmeting eenheid'].value],'\$')
9. Split 89 cell(s) in column Samengevoegd afmeting into several columns by separator

The screenshot shows the 'Undo / Redo' panel with 9 steps. The last two steps are highlighted in blue. The steps are:

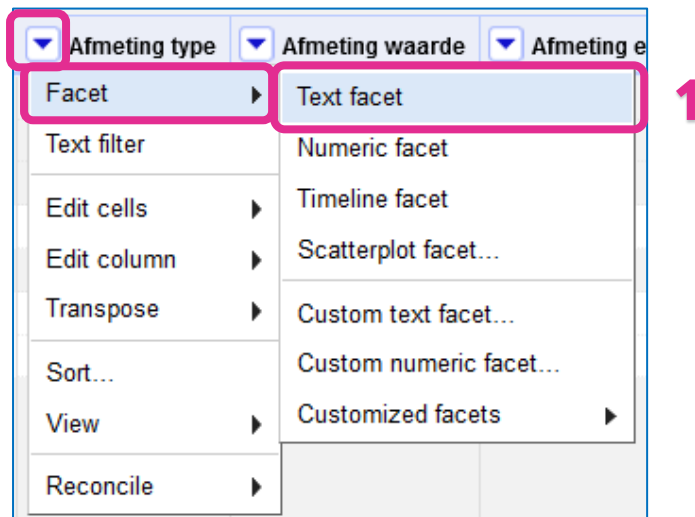
0. Create project
1. Move column Objectnummer to position 0
2. Text transform on 35 cells in column Afmeting type: value.replace(/[\p{Zs}\s]+/, ' ')
3. Split multi-valued cells in column Afmeting type
4. Text transform on 35 cells in column Afmeting waarde: value.replace(/[\p{Zs}\s]+/, ' ')
5. Split multi-valued cells in column Afmeting waarde
6. Text transform on 35 cells in column Afmeting eenheid: value.replace(/[\p{Zs}\s]+/, ' ')
7. Split multi-valued cells in column Afmeting eenheid
8. Create new column Samengevoegd afmeting based on column Afmeting type by filling 149 rows with join ([cells['Afmeting type'].value,cells['Afmeting waarde'].value,cells['Afmeting eenheid'].value],'\$')
9. Split 89 cell(s) in column Samengevoegd afmeting into several columns by separator

Zoeken in OpenRefine

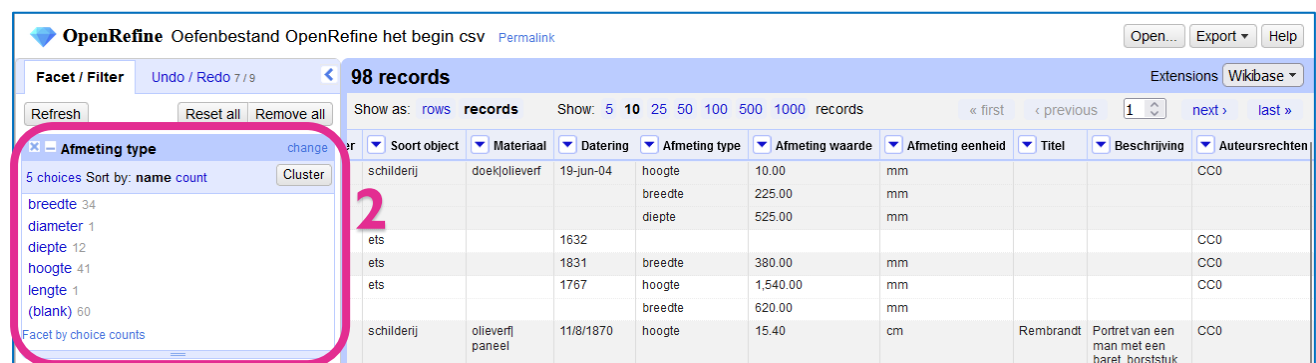
Facet zoeken

Dit is een hele handige manier om overzicht te krijgen in gecontroleerde velden. Het laat namelijk zien welke opties er voorkomen in een bepaalde kolom.

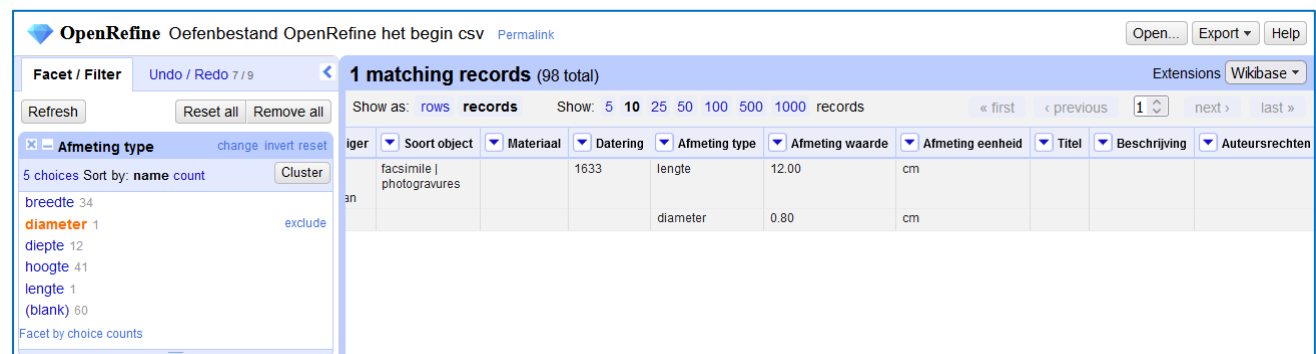
1. Klik op het pijltje voor de kolom die je wilt bevragen, in dit voorbeeld *afmeting type* >
Facet > Tekst facet



2. Er is nu een facetblokje aan de linkerkant van je scherm verschenen



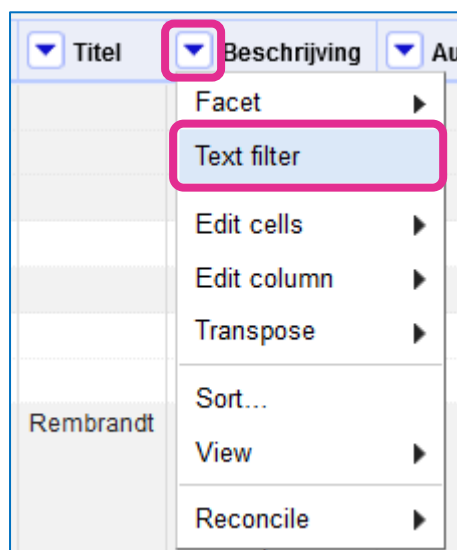
3. Als je nu op één van de waardes klikt, dan verschijnen alleen de records (of rijen) waar deze waarde in voorkomt. In dit voorbeeld is er maar 1 record met "lengte"



Tekst zoeken

Je kunt ook op tekst zoeken in OpenRefine. Dit is bijvoorbeeld erg handig in tekstvelden zoals de beschrijving waar je een specifiek woord in zoekt.

1. Klik op het pijltje voor de kolom die je wilt bevragen > **Tekst filter**



2. Er verschijnt een zoekbalkje links op je scherm



3. OpenRefine laat meteen de records (of rijen) zien waar het gezochte woord in voorkomt.

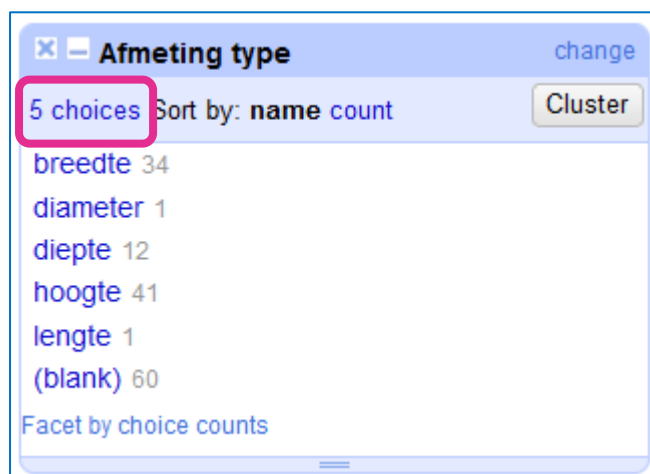
A screenshot of the OpenRefine interface showing the results of the search. The 'Facet / Filter' panel on the left shows the search term 'zelfportret'. The main table displays 2 matching records (98 total). The table has columns for 'rvaardiger', 'Soort object', 'Materiaal', 'Datering', 'Afmeting type', 'Afmeting waarde', 'Afmeting eenheid', 'Titel', 'Beschrijving', and 'Auteurs'. The first record is for a portrait by Rembrandt, and the second is for a self-portrait by Rembrandt van Rijn.

rvaardiger	Soort object	Materiaal	Datering	Afmeting type	Afmeting waarde	Afmeting eenheid	Titel	Beschrijving	Auteurs
randt	schilderij	olieverf paneel	11/8/1870	hoogte	15.40	cm	Rembrandt	Portret van een man met een baret, borststuk en face. Naar een zelfportret van Rembrandt. In bruine lijst met goudkleurige binnenlijst.	CC0
randt			7 augustus 1652	hoogte	34.80		Zelfportret van Rembrandt van Rijn	Zelfportret Rembrandt van Rijn.	CC BY-NC

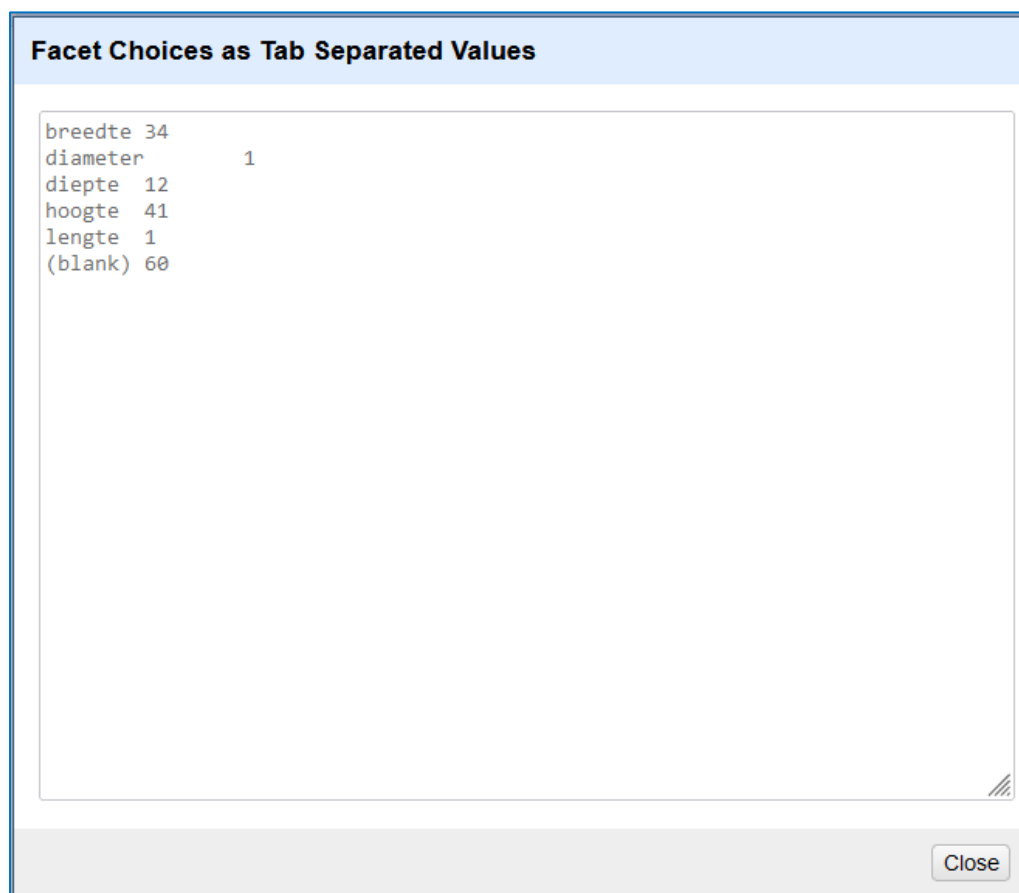
Facet resultaten opslaan

Wil je de lijst met resultaten opslaan zodat je er mee aan de slag kan? Dan kan je de resultaten opslaan in een ander programma, bijvoorbeeld Excel.

1. Klik op **5 choices** (*nb het aantal is het aantal resultaten van het betreffende facet*)



2. Er opent nu een scherm met een TSV indeling van de facet resultaten. Dit kan je selecteren en kopiëren naar een Excel of een ander bestand naar keuze.



Facetten combineren

Je kunt meerdere facetten met elkaar combineren. Zo kan je hele specifieke selecties uitvoeren op je data.

Voorbeeld 1: beschrijving en objectnamen

Als het woord 'rembrandt' in de beschrijving voorkomt, dan zijn dit de vervaardigers die bij de records horen.

The screenshot shows a 'Facet / Filter' panel with the following elements:

- Facet / Filter** header with **Undo / Redo 7 / 9** and a back arrow.
- Buttons: **Refresh**, **Reset all**, and **Remove all**.
- Beschrijving** facet (active):
 - Search input: **rembrandt**
 - Options: ☐ case sensitive, ☐ regular expression
 - Link: **invert reset**
- Vervaardiger** facet (active):
 - Link: **change**
 - Text: **3 choices Sort by: name count**
 - Button: **Cluster**
 - List of results:
 - Rembrandt 4
 - Rembrandt (etser) 1
 - Rembrandt Harmensz. van Rijn (1606-1669) 1
 - (blank) 11
 - Text: **Facet by choice counts**

Voorbeeld 2: afmeting type en afmeting eenheid

Welke eenheden worden gebruikt bij het type 'diepte'?

Je ziet dat met het selecteren van een term bij één facet, de resultaten van het andere facet meteen mee veranderen.

The image shows two side-by-side screenshots of a facet filter interface, illustrating how selecting a term in one facet affects the results of another facet.

Left Screenshot:

- Facet / Filter:** Undo / Redo 7 / 9
- Buttons:** Refresh, Reset all, Remove all
- Facet 1: Afmeting type** (change)
 - 5 choices Sort by: name count Cluster
 - breedte 34
 - diameter 1
 - diepte 12
 - hoogte 41
 - lengte 1
 - (blank) 60
 - Facet by choice counts
- Facet 2: Afmeting eenheid** (change)
 - 4 choices Sort by: name count Cluster
 - 4
 - cm 44
 - m 2
 - mm 25
 - (blank) 74
 - Facet by choice counts

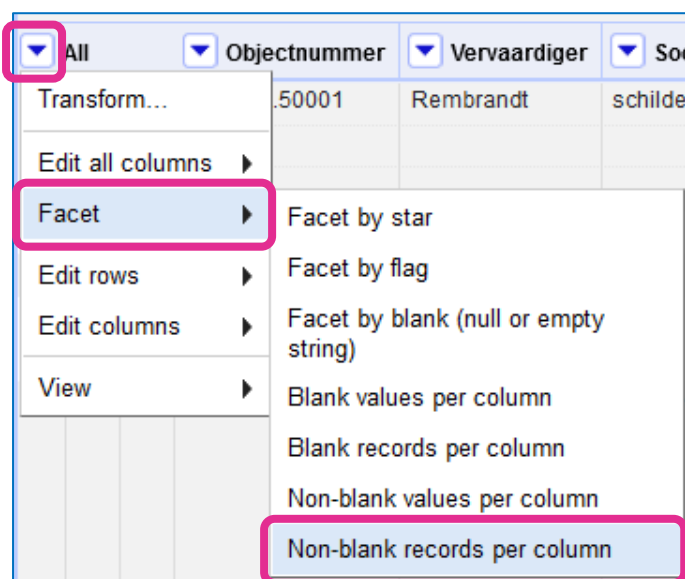
Right Screenshot:

- Facet / Filter:** Undo / Redo 7 / 9
- Buttons:** Refresh, Reset all, Remove all
- Facet 1: Afmeting type** (change invert reset)
 - 5 choices Sort by: name count Cluster
 - breedte 34
 - diameter 1
 - diepte 12 (highlighted in orange, with 'exclude' button)
 - hoogte 41
 - lengte 1
 - (blank) 60
 - Facet by choice counts
- Facet 2: Afmeting eenheid** (change)
 - 2 choices Sort by: name count Cluster
 - cm 10
 - mm 18
 - Facet by choice counts

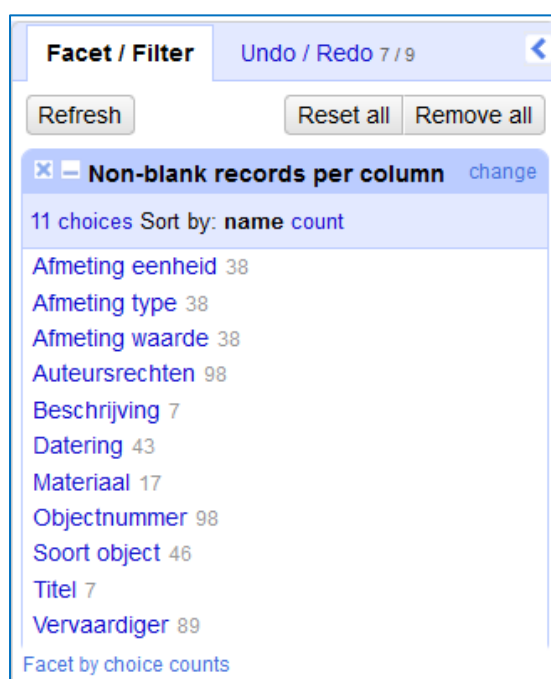
Facet op kolommen

Naast de facetten per kolom (zodat je goed weet wat de inhoud van het veld is), kan je ook facetten uitvoeren over alle kolommen heen. Als je bijvoorbeeld wilt weten welke velden er zijn meegekomen in je export en hoe vaak deze velden gebruikt worden is dit de zoekopdracht die je wilt gebruiken.

1. Klik op het pijltje voor de kolom **All > Facet > Non-blank records per column** (niet-lege records per kolom)



2. Je krijgt nu een overzicht van de velden met een getal erachter wat overeenkomt met de gekozen facet. In dit voorbeeld de ingevulde records per kolom



3. Zo zie je dat het veld "Titel" maar bij 7 objectrecords is ingevuld.

Links en verwijzingen

- OpenRefine user manual: <https://openrefine.org/docs>
- OpenRefine forum: <https://forum.openrefine.org/>
- OpenRefine github: <https://github.com/OpenRefine/>
- Introductie cursus OpenRefine: <https://digitaalerfgoedcoach.online/online-leren-posts/openrefine/>
- Kennisbank Meemoo: <https://kennisbank.meemoo.be/toolbox/open-refine-handleiding-voor-cultureel-erfgoed-collecties>
- Netwerk Digitaal Erfgoed: <https://netwerkdigitaalerfgoed.nl/>
- Datawerkplaatsen: <https://netwerkdigitaalerfgoed.nl/activiteiten/datawerkplaatsen/>
- Landschap Erfgoed Utrecht: <https://www.landschaperfgoedutrecht.nl/>
- <https://www.landschaperfgoedutrecht.nl/erfgoed/digitalisering/>

Alle handleidingen zijn te vinden op: www.kubus.frl/openrefine

