



Data Journalism

Data Cleansing - Open Refine

Clean the data



Definition

Data cleansing, data cleaning or data scrubbing is the process of detecting and correcting (or removing) corrupt or inaccurate records from data.

The term refers to identifying incomplete, incorrect, inaccurate, irrelevant, etc. parts of the data and then replacing, modifying, or deleting this dirty data.

https://en.wikipedia.org/wiki/Data_cleansing

Data cleaning tools



Bibliography

[Open Refine Home page](#)

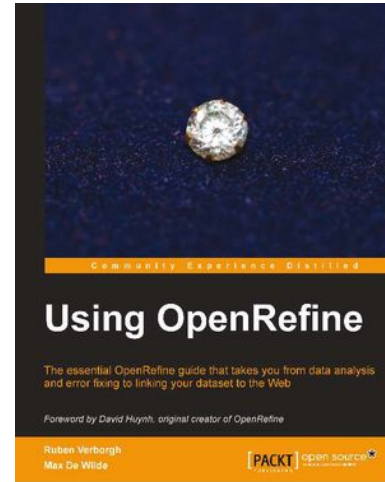
[Official Documentation](#)

[List of Tutorials](#)

[Using OpenRefine Ruben Verborgh, Max De Wilde September 2013](#)

[General Refine Expression Language](#)

[Jython = Python for java platform](#)



What you can do

Cleansing - Analysing and fixing data

- Fixing errors

- Remove duplicate records

- Split multi data columns

Enrichment

- by Web Api

- by Web Scraping

Common Errors you can find in the data

String vs numbers (“10,5432” vs 10.5432)

Different Formats (01/09/2016 vs 01-09-2016)

Data inconsistencies (Piazza, P.zza, P.za)

Lateral spaces (“B&B” vs “ B&B”)

Accommodations in Tuscany

[Regione Toscana - Strutture ricettive](#)



[File CSV](#)



First view with an editor

```
id|nome|lat|lon|codeserc|tipologia|indirizzo|cap|comune|provincia|stelle|email|url|telefono|
14017|FORNI ROSAIA|44.2270033|10.029223799999954|045001AAT0001|Agriturismi|PIAZZA PUCCINI 1
13995|POW WOW DI GRULLI ARISTIDE|44.232061100000003|10.049777499999999|045001AAT0006|Agritur
13989|MONTEBELLO|44.225479999999997|10.029412000000001|045001AAT0011|Agriturismi|Via COLLINA
13820|VALLE FIORITA|44.225616500000001|10.018136999999999|045001AAT0012|Agriturismi|Via AIA
13924|LA SELVA|44.2166706|9.9674972000000004|045001AAT0013|Agriturismi|Selva, 0 - Loc. Selva
13843|FIORENTINI GIANLUCA|44.2166706|9.9674972000000004|045001AAT0014|Agriturismi|sanacco, 1
13832|VILLA MIMOSA|44.174129100000002|9.912286299999999|045001AFR0003|Affittacamere|Via MAE
13783|DEMY|44.215124000000003|9.9673911000000004|045001ALB0002|Alberghi - Hotel|Via Salucci,
13717|PASQUINO|44.2166706|9.9674972000000004|045001ALB0003|Alberghi - Hotel|PIAZZA MAZZINI 2
13848|CASA BARANI|44.205532599999998|9.9698068000000006|045001ALL0003|Alloggi Privati|Via SF
14000|B&B CA' DI MEGOTO|44.225479999999997|10.029412000000001|045001ALL0005|Alloggi Privati|
13803|B&B LO SPIGO|44.232061100000003|10.049777499999999|045001ALL0006|Alloggi Privati|Via M
17369|IL MELOGRANO|44.2166706|9.9674972000000004|045001ALL0010|Alloggi Privati|Libertà, 14/F
21952|B&B CASA RO'|44.174129100000002|9.912286299999999|045001ALL0012|Alloggi Privati|Via A
17871|LE ROCCAGLIE|44.188153999999997|9.939541699999995|045001CAV0003|Case per Vacanze|Saig
14053|GIUNASCO|44.315319600000002|9.9956531000000002|045002AAT0002|Agriturismi|Giunasco - Lo
```

« Start Over

Configure Parsing Options

Project name

9.	13717	PASQUINO	44.2166706	9.9674972000000004	045001ALB0003	Alberghi - Hotel	PIAZZA MAZZINI 22 - Fraz. pippo	54011	Aulla	MS	2	info@albergopasquino.it	www.a
10.	13848	CASA BARANI	44.2055325999999998	9.9698068000000006	045001ALL0003	Alloggi Privati	Via SPRINI 7/A - Fraz. Sprini	54010	Aulla	MS	0	odilbar47@libero.it	www.b
11.	14000	B&B CA' DI MEGOTO	44.2254799999999997	10.029412000000001	045001ALL0005	Alloggi Privati	Via COLLINA 8 - Loc. Olivola	54011	Aulla	MS	0	cadimegato@lunigiana.net	
12.	13803	B&B LO SPIGO	44.2320611000000003	10.0497774999999999	045001ALL0006	Alloggi Privati	Via MONTE BARDOLINI - Loc. Figliolo	54011	Aulla	MS	0	info@lospigo.it	www.lc
13.	17369	IL MELOGRANO	44.2166706	9.9674972000000004	045001ALL0010	Alloggi Privati	LibertÀ , 14/R - Loc. AULLA - Fraz. Albiano Magra	54010	Aulla	MS	0	ilmelogranobeb@hotmail.it	
14.	21952	B&B CASA RO'	44.1741291000000002	9.9122862999999999	045001ALL0012	Alloggi Privati	Via Amola, 18 - Loc. Albiano Magra	54011	Aulla	MS	0	perst16libero.it	
15.	17871	LE ROCCAGLIE	44.1881539999999997	9.9395416999999995	045001CAV0003	Case per Vacanze	Saigola, 8	54011	Aulla	MS	0	info@lerocccaglie.it	www.le
16.	14053	GIUNASCO	44.3153196000000002	9.9956531000000002	045002AAT0002	Agriturismi	Giunasco - Loc. Orturano	54021	Bagnone	MS	2	monique@giunasco.it	www.g
17.	13771	CA' DEL SASSO	44.2881418999999999	9.9982670999999996	045002AAT0005	Agriturismi		54021	Bagnone	MS	2		
18.	13831	FATTORIA DEL DEGLIO	44.3171933	10.013592900000001	045002AAT0006	Agriturismi	Stagnai, 0 - Fraz. Deglio	54021	Bagnone	MS	2	lucy70b@libero.it	
19.	20513	LA VALLE DELLA	44.3146799000000002	9.9945108000000005	045002AAT0007	Agriturismi	Vallescura, 2 - Loc.	54021	Bagnone	MS	2	info@lavalledellalunigiana.it	www.la

Parse data as

Character encoding

CSV / TSV / separator-based files

Line-based text files

Fixed-width field text files

PC-Axis text files

JSON files

MARC files

RDF/N3 files

Wikitext

XML files

Columns are separated by

- commas (CSV)
- tabs (TSV)
- custom |

Escape special characters with \

- Ignore first 0 line(s) at beginning of file
- Parse next 1 line(s) as column headers
- Discard initial 0 row(s) of data
- Load at most 0 row(s) of data



- Parse cell text into numbers, dates, ...
- Quotation marks are used to enclose cells containing column separators

- Store blank rows
- Store blank cells as nulls
- Store file source (file names, URLs) in each row

13397 rows

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

All	id	nome	lat	lon	codeserc	tipologia	indirizzo	
☆	1.	14017	Facet	44.2270033	10.029223799999954	045001AAT0001	Agriturismi	PIAZZA PUCCINI 1 - Loc. Olivola
☆	2.	13995	Text filter	44.2320611	10.0497775	045001AAT0006	Agriturismi	San Domenico la Cavana, 0 - Loc. Bigliolo
☆	3.	13989	Edit cells	44.22548	10.029412	045001AAT0011	Agriturismi	Via COLLINA 7 - Loc. Olivola
☆	4.	13820	Edit column	44.2256165	10.018137	045001AAT0012	Agriturismi	Via AIA DI BELLONE - Loc. Valenza
☆	5.	13924	Transpose	44.2166706	9.9674972	045001AAT0013	Agriturismi	Selva, 0 - Loc. Selva
☆	6.	13843	Sort...	44.2166706	9.9674972	045001AAT0014	Agriturismi	sanacco, 1 - Loc. Quercia
☆	7.	13832	View	44.1741291	9.9122863	045001AFR0003	Affittacamere	Via MAESTRO FERRARI 7 - Loc. Albiano Magra
☆	8.	13783	Reconcile	44.215124	9.9673911	045001ALB0002	Alberghi - Hotel	Via Salucci, 9
☆	9.	13717	Extract named entities...	44.2166706	9.9674972	045001ALB0003	Alberghi - Hotel	PIAZZA MAZZINI 22 - Fraz. pippo
☆	10.	13848	PASQUINO	44.2055326	9.9698068	045001ALL0003	Alloggi Privati	Via SPRINI 7/A - Fraz. Sprini
☆	11.	14000	CASA BARANI	44.22548	10.029412	045001ALL0005	Alloggi Privati	Via COLLINA 8 - Loc. Olivola
☆	12.	13803	B&B CA' DI MEGOTO	44.2320611	10.0497775	045001ALL0006	Alloggi Privati	Via MONTE BARDELLI - Loc. Bigliolo
☆	13.	17369	B&B LO SPIGO	44.2166706	9.9674972	045001ALL0010	Alloggi Privati	Libertà, 14/F - Loc. AULLA - Fraz. Albiano Magra
☆	14.	21952	IL MELOGRANO	44.1741291	9.9122863	045001ALL0012	Alloggi Privati	Via Amola, 18 - Loc. Albiano Magra
☆	15.	17871	B&B CASA RO'	44.188154	9.9395417	045001CAV0003	Case per Vacanze	Saigola, 8
☆	16.	14053	LE ROCCAGLIE	44.3153196	9.9956531	045002AAT0002	Agriturismi	Giunasco - Loc. Orturano

By clicking on the upside down arrow of the column heading you start working on the data

Facet

The screenshot shows a software interface with a table and a context menu. The table has columns for 'tipologia', 'indirizzo', 'cap', 'comune', and 'provin'. The context menu is open over the 'tipologia' column, listing various actions. The 'Text facet' option is highlighted with a red box.

	tipologia	indirizzo	cap	comune	provin
1					
5					
2					
3					
4			54010	Aulla	MS
3			54011	Aulla	MS

FERRARI 7 -

**Text Facet on
“tipologia”
column**

Text Facet

In italian: sfaccettature

technically is an hystogram



Edit Cells

Common transforms

13397 rows

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

All	id	nome	lat	lon	codeserc	tipologia	indirizzo
☆	1.	14017	44.2270033	10.029223799999954	045001AAT0001	Agriturismi	PIAZZA PUCCI Olivola
☆	2.	13995	44.2320611	10.0497775	045001AAT0006	Agriturismi	San Domenico 0 - Loc. Bigliolo
☆	3.	13989		10.029412	045001AAT0011	Agriturismi	Via COLLINA 7
☆	4.	13820					A DI BEL a
☆	5.	13924					0 - Loc.
☆	6.	13843					o, 1 - Lo
☆	7.	13832					ESTRO Albiano M
☆	8.	13783					ucci, 9
☆	9.	13717	44.2166706	9.9			A MAZZI
☆	10.	13848	44.2055326	9.9			RINI 7/A
☆	11.	14000	44.22548	10.			LLINA 8
☆	12.	13803	44.2320611	10.0			ONTE BA
☆	13.	17369	44.2166706	9.9674972	045001ALL0010	Alloggi Privati	Loc. Bigliolo Libertà, 14/F - I - Fraz. Albiano
☆	14.	21952	44.1741291	9.9122863	045001ALL0012	Alloggi Privati	Via Amola, 18 -

**To title
case**

Transforms

	lon	codeserc	tipologia	indirizzo
33			Agriturismi	PIAZZA PUCCINI 1 - Loc. C
11			Agriturismi	San Domenico la Cavana, Bigliolo
48			Transform...	loc. Olivc
65			Common transforms	NE - Loc
06			Fill down	va
06			Blank down	Quercia
91			Split multi-valued cells...	RRARI 7
24			Join multi-valued cells...	
			Cluster and edit...	
06	9.9073	045001ALB0003	Hotel	FRATELLI MARCONI 22 - Fraz
26	9.96981	045001ALL0003	Alloggi Privati	Via SPRINI 7/A - Fraz. Spri

General Refine Expression Language - GREL

Variables

value = value of current cell

row = number of current row

Functions

split("division character")

round() = round up

Custom text transform on column lat

Expression

`round(value*100000)/100000.0`

`round(value*100000)/100000.0`

Preview

History

Starred

Help

row	value	<code>round(value*100000)/100000.0</code>
1.	44.2270033	44.227
2.	44.2320611	44.23206
3.	44.22548	44.22548
4.	44.2256165	44.22562
5.	44.2166706	44.21667
6.	44.2166706	44.21667
7.	44.1741291	44.17413

On error

- keep original
- set to blank
- store error

Re-transform up to times until no change

Edit columns

▼ indirizzo	▼ cap	▼ comune	▼ pro
Facet ▶	54011	Aulla	MS
Text filter ▶	54011	Aulla	MS
Edit cells ▶	54011	Aulla	MS
Edit column ▶			
Transpose ▶			
Sort...			
View ▶			
Reconcile ▶			
Extract named entities...			
PIAZZA MAZZINI 22 - Fraz. pip			
Via SPRINI 7/A - Fraz. Sprini			
Via COLLINA 8 - Loc. Olivola			
Via MONTE BARDELLI - Loc. E			
Libertà 14/F - Loc. ALLIA - Fraz. Albiano Magra	54010	Aulla	MS

- Facet ▶
- Text filter ▶
- Edit cells ▶
- Edit column ▶
- Transpose ▶
- Sort...
- View ▶
- Reconcile ▶
- Extract named entities...
- PIAZZA MAZZINI 22 - Fraz. pip
- Via SPRINI 7/A - Fraz. Sprini
- Via COLLINA 8 - Loc. Olivola
- Via MONTE BARDELLI - Loc. E
- Libertà 14/F - Loc. ALLIA - Fraz. Albiano Magra

- Split into several columns...
- Add column based on this column...
- Add column by fetching URLs...
- Add columns from reconciled values...
- Rename this column
- Remove this column
- Move column to beginning
- Move column to end
- Move column left
- Move column right

Add column based on column indirizzo

New column name

set to blank store error copy value from original column

Expression

Language General Refine Expression Language (GREL) ▾

`value.split(" ")[0].toLowerCase()`

No syntax error.

value.split(" ")[0].toLowerCase()

Preview

History

Starred

Help

row	value	value.split(" ")[0].toLowercas ...
1.	PIAZZA PUCCINI 1 - Loc. Olivola	piazza
2.	San Domenico la Cavana, 0 - Loc. Bigliolo	san
3.	Via COLLINA 7 - Loc. Olivola	via
4.	Via AIA DI BELLONE - Loc. Valenza	via
5.	Selva, 0 - Loc. Selva	selva,
6.	sanacco, 1 - Loc. Quercia	sanacco,
7.	Via MAESTRO FERRARI 7 - Loc. Albiano Magra	via

OK

Cancel

Clustering of values from facet

Cluster & Edit column "Tipo strada"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. [Find out more ...](#)

Method key collision

Keying Function fingerprint

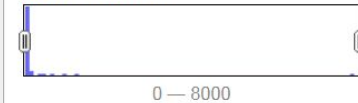
128 clusters found

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
6	7929	<ul style="list-style-type: none">Via (5736 rows)VIA (2154 rows)via (32 rows)Via (5 rows)Via (1 rows)viA (1 rows)	<input type="checkbox"/>	Via
5	82	<ul style="list-style-type: none">Località (72 rows)LOCALITA (3 rows)LOCALITA' (3 rows)Localita' (3 rows)località (1 rows)	<input type="checkbox"/>	Località
5	115	<ul style="list-style-type: none">Lungomare (65 rows)LUNGOMARE (23 rows)Lungomare, (22 rows)lungomare (4 rows)LUNGOMARE, (1 rows)	<input type="checkbox"/>	Lungomare
5	79	<ul style="list-style-type: none">C.S. (66 rows)c.s. (7 rows)C.S., (3 rows)	<input type="checkbox"/>	C.S.

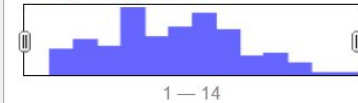
Choices in Cluster



Rows in Cluster



Average Length of Choices



Length Variance of Choices



Select All Unselect All

Export Clusters

Merge Selected & Re-Cluster

Merge Selected & Close

Close

Data enrichment

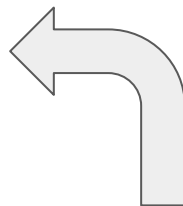
Data Enrichment

Web API - **parseJson(string s)**

Web Scraping - **parseHtml(string s)**

Data Enrichment

Serie	Campionato	Girone	Nome	IndirizzoStadio
b	2017	d	virtus-partenope	Via Sandriana, 80046 San Giorgio a Cremano NA, Italy
b	2017	d	salento-women-soccer	Via Melissano, 73055 Racale LE, Italy
b	2017	d	roma-xiv-decimoquarto	Via di Settebagni, 340, 00139 Roma RM, Italy
b	2017	d	roma-calcio-femminile	Via Verrio Flacco, 41, 00177 Roma RM, Italy
b	2017	d	napoli-calcio-femminile	Via Parroco Giustino Russolillo, 80126 Napoli NA, Italy
b	2017	d	lazio-women-2015	Via Cina, 91, 00144 Roma RM, Italy
b	2017	d	nebrodi	Contrada Vina, 52, 98071 Capo d'Orlando ME, Italy
b	2017	d	latina	Via Tagliamento, 1, 04100 Latina LT, Italy
b	2017	d	grifone-gialloverde	Via Farsalo, 21, 00183 Roma RM, Italy
b	2017	d	catania	Via Mazzola, 95032 Belpasso CT, Italy
b	2017	d	chieti	Viale Abruzzo, 66100 Chieti CH, Italy
b	2017	d	real-colombo	Via Costantino, 5, 00145 Roma RM, Italy
b	2017	d	apulia-trani	Via Superga, 3, 76125 Trani BT, Italy



Lat	Lon
40.832394	14.334764
39.9642814	18.1032099
41.9682803	12.5458501
41.895338	12.5435241
40.8607785	14.1700751
41.8132472	12.4358191
38.136218	14.723183
41.4724754	12.8989634
41.8827569	12.5047246
37.5771625	14.9742431
42.3468217	14.1357738
41.8551649	12.4895363
41.276568	16.402354

Data Enrichment: Open Refine

OpenRefine makes it easy to annotate datasets with data fetched from any web service which returns JSON

I.e. Get coordinates from addresses

need Geocoding Webservice

1. OpenStreetMap
2. GoogleMap

Geocoding services

OpenStreetMap

<http://nominatim.openstreetmap.org/search?format=json&q=Via Moruzzi 1 Pisa>

low recall

GoogleMap

<https://maps.googleapis.com/maps/api/geocode/json?address=Via Moruzzi 1 Pisa&key=YOUR API KEY>

high recall, need a key

Openstreetmap Json Result

```
[  
  {  
    place_id: "16952760",  
    licence: "Data © OpenStreetMap contributors, ODbL 1.0. https://osm.org/copyright",  
    osm_type: "node",  
    osm_id: "1477804118",  
    boundingbox:  
    [  
      "43.7193809",  
      "43.7194809",  
      "10.4237241",  
      "10.4238241"  
    ],  
    lat: "43.7194309",  
    lon: "10.4237741",  
    display_name: "Area della Ricerca del CNR di Pisa, 1, Via Giuseppe Moruzzi, Don Bosco, Pisa,  
    PI, Tuscany, 56124, Italia",  
    class: "place",  
    type: "house",  
    importance: 0.52025  
  }  
]
```

Google Json Result

```
{
  "results":
  [
    {
      "address_components":
      [],
      "formatted_address": "Via Giuseppe Moruzzi, 1, 56127 Pisa PI, Italia" ,
      "geometry":
      {
        "bounds":
        {},
        "location":
        {
          "lat": 43.7182358,
          "lng": 10.4248623
        },
        "location_type": "ROOFTOP",
        "viewport":
        {}
      },
      "place_id": "ChIJ09Em18OR1RIRELeidtGcXhA" ,
      "types":
      []
    }
  ],
  "status": "OK"
}
```

Google API Geocoding service

[Google API geocoding](#) with REST documentation

To use the Google Maps Geocoding API, you need an API key. Before you start developing with the Geocoding API,

review the [authentication requirements](#) and

the [API usage limits](#).

- 2,500 free requests per day (not still available)
- 50 requests per second

[Google Console](#) to manage my API Keys

Data Enrichment with Open Refine

3 rows

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

All	IndirizzoStadio	Serie	Campionato	Giro
1.	gio a	b	2017	d
2.	E, Italy	b	2017	d
3.				

Context menu for the 'IndirizzoStadio' column:

- Facet
- Text filter
- Edit cells
- Edit column**
 - Split into several columns...
 - Add column based on this column...
 - Add column by fetching URLs...**
 - Add columns from Freebase ...
- Transpose
- Sort...
- View
 - Add columns from Freebase ...
- Reconcile
 - Rename this column
 - Remove this column
- Extract named entities...
 - Remove this column
- Move column to beginning
- Move column to end
- Move column left
- Move column right

Data Enrichment with Open Refine

Add column by fetching URLs based on column IndirizzoStadio

New column name Throttle delay milliseconds

On error set to blank store error

Formulate the URLs to fetch:

Expression Language No syntax error.

Preview History Starred Help

row	value	'https://maps.googleapis.com/maps/api/geocode/json?address='+escape(value,'URL')+'&key=YOUR_API_KEY'
1.	Via Sandriana, 80046 San Giorgio a Cremano NA, Italy	https://maps.googleapis.com/maps/api/geocode/json?address=Via+Sandriana%2C+80046+San+Giorgio+a+Cremano+NA%2C+API+KEY
2.	Via Melissano, 73055 Racale LE, Italy	https://maps.googleapis.com/maps/api/geocode/json?address=Via+Melissano%2C+73055+Racale+LE%2C+Italy&key=YOUR_API+KEY
3.	Via di Settebagni, 340, 00139 Roma RM, Italy	https://maps.googleapis.com/maps/api/geocode/json?address=Via+di+Settebagni%2C+340%2C+00139+Roma+RM%2C+Italy+API+KEY

OK Cancel

'https://maps.googleapis.com/maps/api/geocode/json?address='+**escape(value,'URL')**+'&key=YOUR_API_KEY'

[GREL string function](#)

Data Enrichment with Open Refine

The screenshot shows the Open Refine interface. At the top, a dropdown menu is set to 'text'. Below it, a preview of the JSON data is shown, formatted for readability. The data includes a 'formatted_address' field with the value 'Via di Settebagni, 340, 00139 Roma RM, Italy', and a 'geometry' field with 'location' and 'viewport' sub-objects. The 'location' object contains 'lat' (41.9682803) and 'lng' (12.5458501). The 'viewport' object contains 'northeast' and 'southwest' sub-objects with their respective 'lat' and 'lng' values. The 'status' field is 'OK'.

```
{
  "types": [ "postal_code" ]
},
{
  "formatted_address": "Via di Settebagni, 340, 00139 Roma RM, Italy",
  "geometry": {
    "location": {
      "lat": 41.9682803,
      "lng": 12.5458501
    },
    "location_type": "ROOFTOP",
    "viewport": {
      "northeast": {
        "lat": 41.96962928029149,
        "lng": 12.54719990802915
      },
      "southwest": {
        "lat": 41.9669313197085,
        "lng": 12.5445011197085
      }
    }
  },
  "partial_match": true,
  "place_id": "ChIJ04ipcDpkLxMRck--7tSAJfk",
  "types": [ "street_address" ]
},
{
  "status": "OK"
}
}
```

Buttons at the bottom: Apply (Enter), Apply to All Identical Cells (Ctrl-Enter), Cancel (Esc).

If you try to **edit** this cell you can see the formatted json

Data Enrichment with Open Refine

All	IndirizzoStadio	Json	Serie
1.	Via Sandriana, 80046 San Giorgio a Cremano NA, Italy	<pre>Facet exceeded your daily request quota registering for a key at the Google Text filter Edit cells Edit column Split into several columns... Transpose Add column based on this column... Sort... Add column by fetching URLs... View Add columns from Freebase ... Reconcile Rename this column Extract named entities... Remove this column</pre>	b
2.	Via Melissano, 73055 Racale LE, Italy		
3.	Via di Settebagni, 340, 00139 Roma RM, Italy	<pre>metropolitan area of Roma, short "administrative_area_level_2": "Lazio", "short_name": "Lazio" "administrative_area_level_1", "short_name": "IT", "types": ["long_name": "00139", "short "postal_code"] }], "formatted_ 00139 Roma RM, Italy", "geom 41.9682803, "lng": 12.5458501 }, "location_type": "ROOFTOP", "viewport": { "northeast": { "lat": 41.96962928029149, "lng": 12.5471990802915 }, "southwest": { "lat": 41.9669313197085, "lng": 12.5445011197085 } }, "partial_match": true, "place_id": "ChIJo4ipcDpkLxMRCK-7tSAJfk", "types": ["street_address"] }, "status": "OK" }</pre>	

If you try to **edit** this cell you can see the formatted json

[GREL Other Functions](#)

Data Enrichment with Open Refine

Add column based on column Json

New column name

On error set to blank store error copy value from original column

Expression Language Google Refine Expression Language (GREL) ▾

No syntax error.

Preview History Starred Help

```
3. { "results": [ { "address_components": [ { "long_name": "340", "short_name": "340", "types": [ "street_number" ] }, { "long_name": "Via di Settebagni", "short_name": "Via di Settebagni", "types": [ "route" ] }, { "long_name": "La Cinquina - Bufalotta", "short_name": "La Cinquina - Bufalotta", "types": [ "locality", "political" ] }, { "long_name": "Roma", "short_name": "Roma", "types": [ "administrative_area_level_3", "political" ] }, { "long_name": "Città Metropolitana di Roma", "short_name": "RM", "types": [
```

OK Cancel

`value.parseJson().results[0].geometry.location.lat`

The Google Geocoding Service returns always a list of results[], we get the first: results[0]

Export Data

Open...

Export ▾

Help

Export project

Tab-separated value

Comma-separated value

HTML table

Excel (.xls)

Excel 2007+ (.xlsx)

ODF spreadsheet

Triple loader

MQLWrite

Custom tabular exporter...

Templating...

RDF as RDF/XML

RDF as Turtle

RDF ▾

> last

nt@gr

gopiet

ends.it

erossit

uelune

@gma

ino.it

lia.lu.it

sonbe

ora@g

miamibeb.com

www.miamibeb.com