

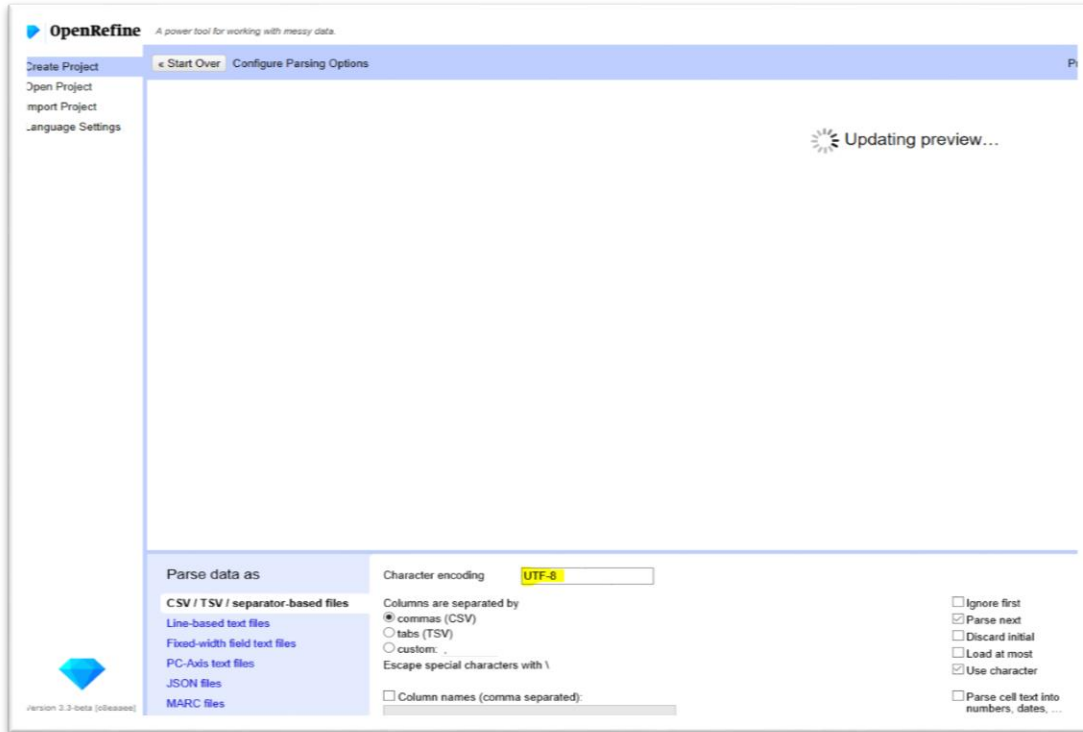


# Data cleaning and validation

# Step1. Uploading BOJA file



# Step2. Giving right format





# Step3. Giving format to date field

(fecha de disposición)

127.0.0.1:3333/project?project=1600054422847

Info NEWSKIES Importado de boquer

OpenRefine boja.csv Personalize

Facet / Filter Undo / Redo 2 / 2

50015 records

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

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](#)

	Organización	Tipo de disposic	Número de dispo	Número del bolet	Fecha del bolet	Fecha de la disp	Sección	Subsecc
gto de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos actos sancionatorios en materia de legislación medioambiental.	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	75	150	07/08/2017	2017-08-02T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales
gto de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos actos sancionatorios en materia de legislación medioambiental.	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	74	150	07/08/2017	2017-08-02T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales
gto de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos actos sancionatorios en materia de legislación medioambiental.	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	73	150	07/08/2017	2017-08-02T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales
mayo de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Jaén, por el que se da fe Ambiental Estratégica de la Identificación Particular Plan Especial de Reforma Interior de la Unidad de Ejecución P-3, sal de Luque (Jaén) (PP 1437/2017).	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	72	150	07/08/2017	2017-05-18T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales
junio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Cádiz, por el que se abre información pública y se dispone la publicación del Proyecto Básico de Mejora de Accesibilidad a la Playa del Castillo en Alca, como paso previo a la otorgación de título concesional para la ocupación de terrenos del dominio público marítimo-terrestre (R 2017).	Consejería de Medio Ambiente y Ordenación del Territorio	Acuerdos	71	150	07/08/2017	2017-06-16T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales
junio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Almería, por el que se y información pública del expediente de autorización para realizar obras en zona de policía del cauce Ramba de riego municipal de Dalmeida (PP 2017/2017).	Consejería de Medio Ambiente y Ordenación del Territorio	Acuerdos	70	150	07/08/2017	2017-06-27T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales
de junio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Sevilla, por la que se resuelve la subvención económica con carácter de subvención para el proyecto de mejora del sistema de gestión de residuos.	Consejería de Medio Ambiente y Ordenación del Territorio	Resoluciones	69	150	07/08/2017	2017-06-16T00:00:00Z	5. Anuncios	5.2. Otros anuncios oficiales

# Step4. TimeLine Facet (Fecha de disposición)

OpenRefine boja.csv

Facet / Filter Undo / Redo 2 / 2

Refresh Reset All Remove All

500515 records

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

Fecha de la disposición change facet

12-83-28 02:00:00 — 5291-83-06 01:00:00

☒ Time ☐ Non-Time ☐ Blank ☐ Error

	Organización	Tipo de diagnóstico	Número de diagnóstico	Número del expediente	Fecha del dictamen	Fecha de la disposición	Sección	Subsección
grito de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos actos sancionadores en materia de legislación medioambiental.	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	73	150	07/09/2017	2017-06-02700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
grito de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos actos sancionadores en materia de legislación medioambiental.	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	74	150	07/09/2017	2017-06-02700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
grito de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos actos sancionadores en materia de legislación medioambiental.	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	73	150	07/09/2017	2017-06-02700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
mayo de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Jaén, por el que se da fe Ambiental Explotación de la Modificación Puntual Plan Especial de Reforma Interior de la Unidad de Ejecución P-3-Jaén de Lineros (Jaén), (PP. 1437/2017).	Consejería de Medio Ambiente y Ordenación del Territorio	Anuncios	72	150	07/09/2017	2017-05-18700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
junio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Cádiz, por el que se abre información pública y se dispone la publicación del Proyecto Básico de Maqueta de Accesibilidad a la Playa del Castillo en ABC, como parte previa a la obtención de Urbanización de la zona de la playa del Castillo marítimo (Cádiz).	Consejería de Medio Ambiente y Ordenación del Territorio	Acuerdos	71	150	07/09/2017	2017-06-16700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
junio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Huelva, por la que se informa pública del expediente de autorización para realizar obras en zona de policía del cauce Riancho de riego municipal de Badajoz (PP. 2017/2017).	Consejería de Medio Ambiente y Ordenación del Territorio	Acuerdos	70	150	07/09/2017	2017-06-27700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
de junio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Sevilla, por la que se resuelve de autorización ambiental unificada para el proyecto que se ubica, en el término municipal de Écija, provincia (Sevilla).	Consejería de Medio Ambiente y Ordenación del Territorio	Resoluciones	69	150	07/09/2017	2017-06-26700-00-002	5. Anuncios	5.2. Otros anuncios oficiales
de julio de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Cádiz, por la que se informa pública del expediente de autorización para realizar obras en zona de policía del cauce Riancho de riego municipal de Badajoz (PP. 2017/2017).	Consejería de Medio Ambiente y Ordenación del Territorio	Resoluciones	68	150	07/09/2017	2017-07-19700-00-002	5. Anuncios	5.2. Otros anuncios oficiales

# Giving format to date fields and Date facets

(fecha de publicación & fecha del boletín)

The screenshot shows the OpenRefine web interface with a dataset of 49,815 records. The left sidebar displays three date faceting panels: 'Fecha de la disposición', 'Fecha de publicación', and 'Fecha del boletín'. Each panel shows a histogram of dates and a list of selected date ranges. The main table displays the data with columns for various fields, including dates. The 'Fecha del boletín' column shows dates in a standard format (YYYY-MM-DD). The 'Fecha de publicación' column shows dates in a standard format (YYYY-MM-DD). The 'Fecha de la disposición' column shows dates in a standard format (YYYY-MM-DD). The table also includes columns for 'Organización', 'Tipo de disposición', 'Número de días', 'Número del boletín', 'Sección', 'Subsección', and 'Fecha del país'.

Fecha de la disposición	Organización	Tipo de disposición	Número de días	Número del boletín	Fecha del boletín	Fecha de la disp.	Sección	Subsección	Fecha del país	
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	75	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	74	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	73	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	72	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	71	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	70	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	69	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	68	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	67	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	66	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00
1912-03-28 19:00:00 — 2301-03-06 19:00:00	Agente de 2017, de la Delegación Territorial de Medio Ambiente y Ordenación del Territorio en Málaga, sobre diversos asuntos sancionadores en materia de legislación medioambiental.	Comunicación de Medio Ambiente y Ordenación del Territorio	Anuncios	65	150	2017-07-06T00:00:00	2017-08-02T00:00:00	5. Anuncios	5.2. Otros anuncios oficiales	2017-08-02T00:00:00



# Step5. Selecting highers values depending on:

Fecha de la disposición + Fecha de publicación + Fecha del boletín

OpenRefine interface showing a table of 68224 matching records. The table displays various records with columns for Organización, Tipo de disposic, Número de diag, Número del bolet, Fecha del bolet, Fecha de la diag, and Sección. The interface includes filters for Fecha de la disposición, Fecha de publicación, and Fecha del boletín.

Organización	Tipo de disposic	Número de diag	Número del bolet	Fecha del bolet	Fecha de la diag	Sección
Comisión de Agricultura y Pesca	Órdenes	1	153	2009-07-01	2009-07-01	1
Comisión de Educación	Órdenes	2	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	3	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	4	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	5	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	6	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	7	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	8	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	9	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	10	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	11	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	12	153	2009-07-01	2009-07-01	2
Comisión de Salud	Resoluciones	13	153	2009-07-01	2009-07-01	2

# Step 6. Text Facet and CLUSTER

(Tipo disposición)

**Cluster & Edit column "Tipo de disposición"**

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 **nearest neighbor** **levenshtein** Radius **1.0** Block Chars **6** **17 clusters found**

		<ul style="list-style-type: none"><li>• <b>notificaciones</b> (16 rows)</li></ul>		
2	2	<ul style="list-style-type: none"><li>• <b>Corrección de Errores</b> (1 rows)</li><li>• <b>Corrección de errores</b> (1 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Corrección de Errores"/>
2	349	<ul style="list-style-type: none"><li>• <b>Decreto del Presidente</b> (251 rows)</li><li>• <b>Decretos del Presidente</b> (98 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Decreto del Presidente"/>
2	269	<ul style="list-style-type: none"><li>• <b>Correcciones de Erratas</b> (167 rows)</li><li>• <b>Correcciones de erratas</b> (102 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Correcciones de Erratas"/>
2	59	<ul style="list-style-type: none"><li>• <b>Decretos-leyes</b> (44 rows)</li><li>• <b>Decretos Leyes</b> (15 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Decretos-leyes"/>
2	10378	<ul style="list-style-type: none"><li>• <b>Órdenes</b> (10370 rows)</li><li>• <b>Ordenes</b> (8 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Órdenes"/>
2	16499	<ul style="list-style-type: none"><li>• <b>Acuerdos</b> (10770 rows)</li><li>• <b>Acuerdo</b> (5729 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Acuerdos"/>
2	149968	<ul style="list-style-type: none"><li>• <b>Anuncios</b> (87691 rows)</li><li>• <b>Anuncio</b> (62277 rows)</li></ul>	<input type="checkbox"/>	<input type="text" value="Anuncios"/>

**# 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



## Cluster & Edit column "Tipo de disposición"

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...](#)

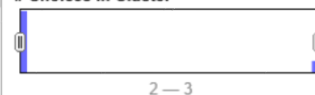
Method key collision

Keying Function fingerprint

6 clusters found

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
3	6951	<ul style="list-style-type: none"> <li>CORRECCION DE ERRORES (6908 rows)</li> <li>Corrección de errores (37 rows)</li> <li>Corrección de Errores (6 rows)</li> </ul>	<input checked="" type="checkbox"/>	CORRECCION DE ERRORES
2	2	<ul style="list-style-type: none"> <li>Corrección de Errores (1 rows)</li> <li>Corrección de errores (1 rows)</li> </ul>	<input checked="" type="checkbox"/>	Corrección de Errores
2	2200	<ul style="list-style-type: none"> <li>Correcciones de Errores (1185 rows)</li> <li>Correcciones de errores (1015 rows)</li> </ul>	<input checked="" type="checkbox"/>	Correcciones de Errores
2	4002	<ul style="list-style-type: none"> <li>Notificaciones (3986 rows)</li> <li>notificaciones (16 rows)</li> </ul>	<input type="checkbox"/>	Notificaciones
2	269	<ul style="list-style-type: none"> <li>Correcciones de Erratas (167 rows)</li> <li>Correcciones de erratas (102 rows)</li> </ul>	<input type="checkbox"/>	Correcciones de Erratas
2	10378	<ul style="list-style-type: none"> <li>Órdenes (10370 rows)</li> <li>Ordenes (8 rows)</li> </ul>	<input type="checkbox"/>	Órdenes

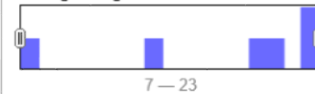
# Choices in Cluster



# Rows in Cluster



Average Length of Choices



Select All Unselect All

Export Clusters

Merge Selected & Re-Cluster

Merge Selected & Close

Close