

# Pattern matching with Regular Expressions


These websites provides a very useful list of symbols and syntax used in regular expressions, and contains examples and exercises.

<https://librarycarpentry.org/lc-data-intro/01-regular-expressions.html>

Tutorial: <https://regex101.com/>

**Case example of using regular expressions** to extract a component of a field that can't be parsed out via a json extract call (thank you to Jean Pajerek for this):

The "value" field of the configuration\_entries table contains many extraneous characters. In this example, we need to extract the "name" component:

configuration_entries  Enter a SQL expression to filter results (use Ctrl+Space)					
	_id	metadata_updated_date	module	value	
34	73128de	2021-06-16 17:40:46.573 -0400	USERSBL	/reset-password	
35	73128de	2021-06-17 17:26:35.408 -0400	TENANT	{"name":"LTS Acquisitions","address":"LTS Acquisitions\n110 C	
36	73128de	2021-06-17 17:26:46.160 -0400	TENANT	{"name":"LTS Approvals","address":"LTS Approvals\n110 C	
37	73128de	2021-06-17 17:27:01.360 -0400	TENANT	{"name":"LTS Documents","address":"LTS Documents\n110 C	
38	73128de	2021-06-17 17:27:12.355 -0400	TENANT	{"name":"LTS E-Resources & Serials","address":"LTS E-Resi	
39	73128de	2021-06-17 17:27:24.618 -0400	TENANT	{"name":"LTS Standing Orders","address":"LTS Standing O	

**SELECT**

cfge.id AS bill\_to\_id,

substring (cfge.value, '([A-Z].+)(", ".+)' ) as bill\_to\_name

**FROM**

configuration\_entries AS cfge

**WHERE** cfge.value LIKE '{"name"%'

- This regular expression says: select the substring that meets the following pattern:
  - the first component starts with any letter and is followed by any number of other characters (the dot is a wildcard and the plus means any number of them);
  - The second component begins with the character sequence "," and is followed by any number of other characters
  - the Where clause stipulates that the "value" field must begin with the character sequence '{"name"%'