

How to use LETRAS in Web

ver. 2014/7/25


LETRAS-web for advanced search

Language: **English** ▾

Output page: **This page** ▾

Explanation of file **Select** ▾

How to use [PDF]



2014-07-31 / 03:57:43

>Top page

(ver. 2014.7.25)

EXECUTE

[1] Input

(a) File: **CODEA-cf** ▾

(b) Filter: **All** ▾
#CODEA
Tipo
#Año
Lugar

(c) Item: **Rg** ▾
AN
CV
CN
Portugal

[2] Output

(a) **7** Sequence table ▾ / (b) Max. output lines: **1000** ▾ / (c) Sort

[3] Pattern

`%n<n>%`
`%<n>n%`
`%nn%`
`%ñ%`

Lower-upper case / distinction

For copy: áéíóúúÁÉÍÓÚÚßçðñÑñrτził

(a) Regular expression (¥ = \):

`[a1-a2], (...), (r1|r2), ., ?, *, +, {n1, n2}, ¥, ^, $, ¥w/¥W, ¥d/¥D, ¥b/¥B, ¥s/¥S, ¥r, ¥n, ¥t`

(b) Expanded regular expression in LETRAS (¥ = \):

`#, %, &, @, ~, ¥l/¥L, ¥v/¥V, ¥c/¥C, ¥h/¥H, ¥k/¥K, ¥z/¥Z, ¥j/¥J, ¥y/¥Y, ¥e/¥E, ¥g/¥G ¥i/¥I, r1+<R>+r2, r1-<R>-r2`

[4] Frequency table / Sequence table

(a) DIVISION: **Total** ▾
Pattern
Sequence

(b) CLASS: **#Año** ▾
Lugar
Provincia

Sort Sort

(c) Interval: **50** ▾
100
200

(d) Calculation: **Horizontal %** ▾
Per mil words
Per mil letters

Hiroto Ueda, University of Tokyo

[0] To begin with

Language:
 ▾

Output page:
 ▾

Explanation of file
 ▾

[How to use \[PDF\]](#)

Select language, output page.

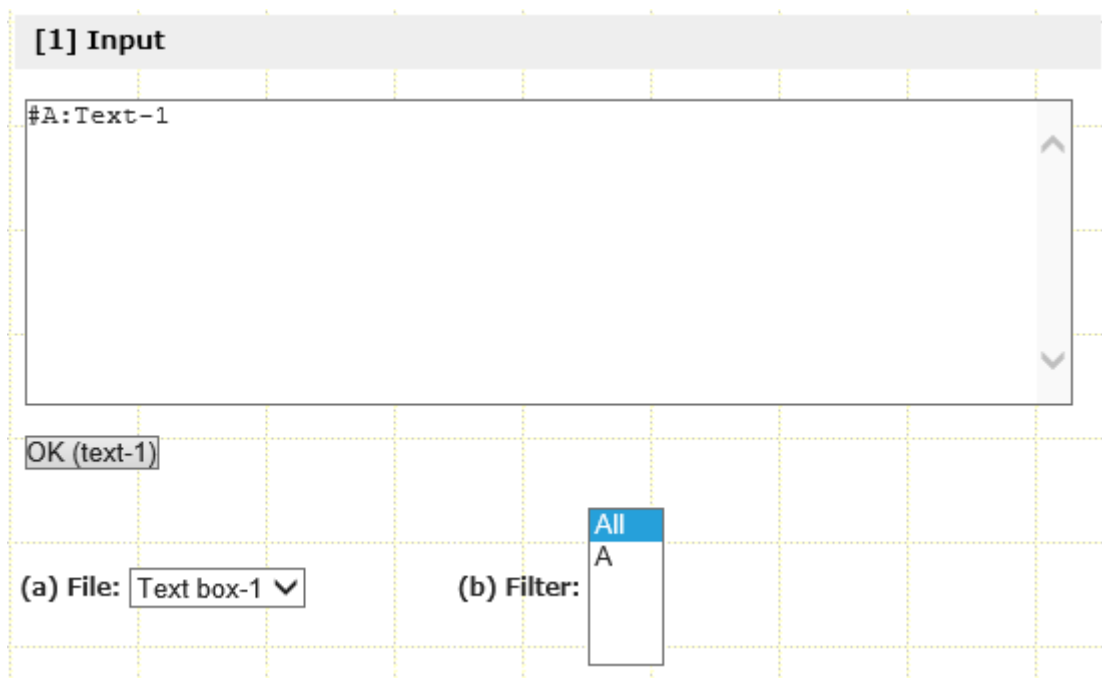
1. Input

[1] Input

(a) File: ▾

(b) Filter: ▾
ID
LUGAR
TIPO
NIVEL

In [1] Input, we select one of the items listed on the pull-down. When selecting Text box the following text box appears:



The first line is for title. In the next line, paste the text data, object of the analysis. When selecting the «Filter» out of «All» (multiple selection is possible), the «Items» box will be opened. Select the items (multiple selection is possible):

In text box, you can set meta-data in this form: #...:..., for example:

#Place:Madrid

#Age:40

En el centro de Madrid se tropiezan [...]

- ¡Hombre, José! ¿Qué tal? ¿Cómo te va? Que vas acelerado, para un poco, oye, ¿tú sabes dónde está ***?

#Age:30

- Hombre, claro. Mira, tiras a la derecha [por la calle Burgos [...]

#Age:40

- Para, para, Toñín, que no quiero ir en coche. Mejor será que coja el bus, ¿no?

#Place:Lima

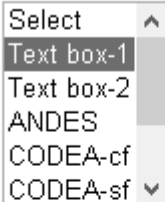
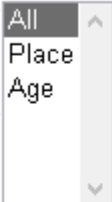
#Age:40

- Me lo hubieras dicho antes, pues. No hay problema. Acá en la esquina puedes tomar un micro, [...]

#Age:30

- Ya, pero ¿cómo sé dónde me bajo? Porque los microbuseros son unos pesados . [...]


[1] Input


(a) File:  (b) Filter: 

```
#Place:Madrid
#Age:40
En el centro de Madrid se tropiezan [...]
- ¡Hombre, José! ¿Qué tal? ¿Cómo te va? Que vas acelerado, para
un poco, oye, ¿tú sabes dónde está ***?
#Age:30
- Hombre, claro. Mira, tiras a la derecha [por la calle Burgos
[...]]
#Age:40
- Para, para, Toñín, que no quiero ir en coche. Mejor será que
```

Server's file:

[1] Input

File name: Filter: 

Item: 

2. Output

[2] Output

(a) / (b) Max. output lines: / (c) Sort

Select (a) output mode, (b) Max. output lines and (c) Sort. In output mode, there are: 0. Sequence, 1. Sequence in all text, 2. Sequence in text, 3. Sequence out of text, 4. Sequence in context, 5. Input text, 6. Frequency table and 7. Sequence frequency table.

3. Pattern

[3] Pattern

Lower-upper case / distinction

For copy: áéíóúüÁÉÍÓÚÜßçðñÑñrτζιζ

(a) Regular expression (¥ = \):

[a1-a2], (...), (r1|r2), ., ?, *, +, {n1, n2}, ¥, ^, \$, ¥w/¥W, ¥d/¥D, ¥b/¥B, ¥s/¥S, ¥r, ¥n, ¥t

(b) Expanded regular expression in LETRAS (¥ = \):

#, %, &, @, ~, ¥l/¥L, ¥v/¥V, ¥c/¥C, ¥h/¥H, ¥k/¥K, ¥z/¥Z, ¥j/¥J, ¥y/¥Y, ¥e/¥E, ¥g/¥G ¥i/¥I, r1+<R>+r2, r1-<R>-r2

Specify the search patten and select Lower-upper case distincion. In the search patten, we can use the following regular expressions.

(a) Reguar expressions:

Bracket [a1-a2]	Letter class, from a1 to a2
------------------	-----------------------------

Parenthesis ()	Subpattern
r1 r2	Selection
Period (.)	All letter excepto carriage return.
Question mark (?)	zero or one of the preceding element, minimum repetition
Asterisk (*)	zero or more of the preceding element
Plus (+)	one or more of the preceding element
{n1, n2}	Matches the preceding element at least n1 and not more than n2 times
¥, back slash	Escape
Caret (^)	Starting position, exception in letter class []
Dollar (\$)	Endin position
¥w, ¥W	Alphanumeric character, including "_";, its complement
¥d, ¥D	Digit [0-9], its complement
¥b, ¥B	English word boundary, its complement
¥s, ¥S	Tab, line feed, form feed, carriage return, and space, its complement
¥r	Line feed
¥n	Carriage return
¥t	Tab

(b) Expanded reguar expressions in LETRAS:

#	Word boundary
%	Zero or more letters
&	One or more letters
@, ~	One letter
¥l, ¥L	Expanded latin letter, its complement
¥v, ¥V	Vocalic letter, its complement
¥c, ¥C	Consonantal letter, its complement
¥h, ¥H	Hiragana, its complement
¥k, ¥K	Katakana, its complement
¥z, ¥Z	Kanzi, its complement

¥j, ¥J	Japanese letter, its complement
¥i, ¥I	Cyrillic letter, its complement
¥g, ¥G	Greek letter, its complement
¥y, ¥Y	Expanded Chinese letter, its complement
¥e, ¥E	Korean letter, its complement
r1<<r2>>r3	Positive context
r1!<r2>!r3	Negative context
n@	Sequence of <i>n</i> letters, for example 3@
n&	Sequence of <i>n</i> words, for example 3&

4. Frequency table / Sequence table

[4] Frequency table / Sequence table

(a) DIVISION: Total
Pattern
Sequence

Sort

(b) CLASS: All
ID
LUGAR

Sort

(c) Interval: 0
10
20

(d) Calculation: Absolute frequency
Vertical %
Horizontal %

When selecting [6. Frequency table] or [7. Sequence table] in [2] Output, we specify (a) Division, (b) Class, (c) Interval and (d) Calculation. In (a) Division are Total, Pattern, Sequence; in (b) Class are titles of Input data, in (c) Interval, interval quantity, in (d) Calculation, Absolute frequency, vertical relative frequency (%), horizontal relative frequency (%), Per mil words and per mil letters.

5. Examples of output

(1) [Sequence] Pattern: el

Sequence					
File: VARITEX / Execution time: 0.182 s. /					
Output lines: 905 / Max. output lines: 1000 / Sort: ON / Pattern[1]: el; Lower-upper case distinction: ON					
#Op	Sequence	CIUDAD	TEMA	Línea	#Ip.
1	el	Madrid	(A)_Hotel	1	1
2	el	Madrid	(A)_Hotel	1	1
3	el	Madrid	(A)_Hotel	1	1
4	el	Madrid	(A)_Hotel	1	1

(2) [Sequence in all text] Patten: el; Lower/upper case distinction (ON)

Sequence in all text					
File: VARITEX / Execution time: 0.020 s. /					
Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: el; Lower-upper case distinction: ON					
#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(3) [Sequence in all text] Patten: el; Lower/upper case distinction (OFF)

Sequence in all text					
File: VARITEX / Execution time: 0.021 s. /					
Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: el; Lower-upper case distinction: OFF					
#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(4) #el# (#:Word boundary)

Sequence in all text

File: VARITEX / Execution time: 0.053 s. /

Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #el#; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(5) #(el|la|los|las)# (** | **): Subpattern

Sequence in all text

File: VARITEX / Execution time: 0.044 s. /

Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #(el|la|los|las)#; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(6) #una?# (: Repetition from 0 to 1)

Sequence in all text

File: VARITEX / Execution time: 0.019 s. /

Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #una?#; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(7) #p% (Word beginning with 'p')

Sequence in all text

File: VARITEX / Execution time: 0.021 s. /

Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #p%; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(8) #Yv% (Word beginning with vowel)

Sequence in all text

File: VARITEX / Execution time: 0.054 s. /

Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #Yv%; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial . El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(9) #Yv%Yc# (Word beginning with vowel and ending with consonant)

Sequence in all text

File: VARITEX / Execution time: 0.050 s. /

Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #Yv%Yc#; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial . El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(10) &s& (word with 's' inside)

Sequence in all text

File: VARITEX / Execution time: 0.050 s. /
 Output lines: 1001 / Max. output lines: 1000 / Pattern[1]: #\v%#; Lower-upper case distinction: OFF

#Op	TEXTO	CIUDAD	TEMA	Línea	#Ip.
1	A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	Madrid	(A)_Hotel	1	1
	- ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio				

(11) Frequency table

#auto(bús|buses)#

#micros?#

#cami(ón|ones)#

Frequency table

File: VARITEX / Execution time: 0.110 s. /
 Output lines: 29 / Max. output lines: 1000 / Sort: ON / Pattern[1]: #auto(bús|buses)#; [2]: #micros?#; [3]: #cami(ón|ones)#; Lower-upper case distinction: OFF
 DIVISION: Pattern / Sort: ON / CLASS: CIUDAD / Sort: ON // Interval: 0 / Calculation: Absolute frequency

#Op	Pattern	B.A.	Lima	Madrid	México	Sevilla
1	#auto(bús buses)#			4		10
2	#cami(ón ones)#				2	
3	#micros?#	3	6		3	

(12) Sequence table

Sequence table

File: VARITEX / Execution time: 0.137 s. /
 Output lines: 28 / Max. output lines: 1000 / Sort: ON / Pattern[1]: #auto(bús|buses)#; [2]: #micros?#; [3]: #cami(ón|ones)#; Lower-upper case distinction: ON
 DIVISION: Pattern / Sort: ON / CLASS: CIUDAD / Sort: OFF // Interval: 0 / Calculation: Absolute frequency

#Op	[Pattern]	Madrid	Sevilla	México	Lima	B.A.
1	#auto(bús buses)#	autobús-102	autobús-114			
2	#auto(bús buses)#	autobús-104	autobús-116			
3	#auto(bús buses)#	autobús-110	autobús-117			
4	#auto(bús buses)#		autobús-124			
5	#auto(bús buses)#		autobús-183			

(13) Frequency table (Class = Año / Interval =50 / Per mil words)

%<n>n%

%<n>n%

%nn%

Frequency table

File: CODEA-cf / Execution time: 5.166 s. / Filter : [1] Región = (1) CV; (2) CN;
 Output lines: 4918 / Max. output lines: 1000 / Sort: ON / Pattern:[1]: %<n>n%; [2]: %<n>n%; [3]: %nn%; Lower-upper case
 distinction: ON
 DIVISION: Pattern / Sort: ON / CLASS: #Año / Sort: ON // Interval: 50 / Calculation: Per mil words

#Op	\$Pattern	1100	1150	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650
3	%<n>n%			1.711	0.248	0.035	0.018	0.015		0.010	0.017		
1	%<n>n%			2.688	8.218	8.216	7.261	9.674	9.437	10.161	5.890	4.825	2.522
2	%nn%	10.020		4.888	2.057	0.628	0.111	0.077	0.394	0.253	0.871	1.076	0.039

(14) Sequence table (Class = Año / Interval =50)

Sequence table

File: CODEA-cf / Execution time: 5.559 s. / Filter : [1] Región = (1) CV; (2) CN;
 Output lines: 4918 / Max. output lines: 1000 / Sort: ON / Pattern:[1]: %<n>n%; [2]: %<n>n%; [3]: %nn%; Lower-upper case
 distinction: ON
 DIVISION: Pattern / Sort: ON / CLASS: #Año / Sort: ON // Interval: 50 / Calculation: Per mil words

#Op	\$Pattern	1100	1150	1200	1250	1300	1350	1400
1	%<n>n%			do<n>nabela-11034	A<n>nos-2324	se<n>nalado-3789	co<n>nusco-21653	co<n>n
2	%<n>n%			Co<n>noscida-39144	Se<n>nor-2327	co<n>nosçemos-4297		
3	%<n>n%			ye<n>neguez-39145	vi<n>nas-2328			
4	%<n>n%			Co<n>noscida-39147	Se<n>nor-2331			
5	%			ye<n>neguez-	Se<n>nor-2333			

FIN