Datos: lowerbeginner.txt

1. ¿Cuántas líneas, palabras y caracteres tienen tus composiciones?

* 10826 palabras en LowerBeginner.txt
* 92 líneas LowerBeginner.txt
* 60716 caracteres en LowerBeginner.txt

1. ¿Qué hace la siguiente línea de código? Explica la función de cada comando.

$ sed 's/[^a-zA-Z]\+/\n/g' < LowerBeginner.txt | sort | uniq -c | sort -nr | sed 10q

sed 's/[^a-zA-Z]\+/\n/g' < LowerBeginner.txt | sort | uniq -c | sor

t -nr | sed 20q

1. ¿Cuáles son las 20 palabras más frecuentes del corpus?
2. 377 es
3. 374 y
4. 288 en
5. 258 a
6. 248 de
7. 233 el
8. 201 la
9. 172 Yo
10. 166 muy
11. 155 gusta
12. 150 un
13. 135 que
14. 108 mi
15. 99 Mi
16. 95
17. 89 El
18. 88 Ella
19. 87 Me
20. 83 no
21. 78 Le

4. Escoge cualquier tipo de afijo gramátical o léxico que te interese. ¿Cuántas veces lo utilizaste?

tr ' ' '\n' < muestra.txt | grep -c '\<in'

60

localhost:~$ tr ' ' '\n' < muestra.txt | grep -c 'ción\>'

92

localhost:~$ tr ' ' '\n' < muestra.txt | grep -c 'sión\>'

25

localhost:~$ tr ' ' '\n' < muestra.txt | grep -c 'sion\>'

0

localhost:~$ tr ' ' '\n' < muestra.txt | grep -c 'cion\>'

0

Pantallazo