

```
tr -sc 'A-Za-z' '\012' < datafile | sort | uniq -c > output
```

計算行數，字數，字元數的程式 wc

計算行數 `wc -l datafile`

計算字數 `wc -w datafile`

計算字元數 `wc -c datafile`

四、從檔案中將包含某一字串或形式的行列出來 grep

從檔案中將包含 pattern 這個字串的行列出來 `grep 'pattern' datafile > output`

五、awk 一種簡單但功能強大的程式語言

從檔案中將包含 pattern 這個字串的行列出來

`awk '/pattern/ { print }' datafile > output`

列印 datafile 第二欄與第一欄中間以 tab 間隔

`awk '{ print $2 "\t" $1 }' datafile > output`

如果 datafile 第三欄的值大於 1.65 列印第二欄與第三欄，中間並以 tab 間隔

`awk '{ if ($3 >= 1.65) { print $2 "\t" $3 } }' datafile > output`

六、perl：結合了 C 語言，awk，sed，shell programing，功能比 awk 更強大，

非常適合處理語料。