# Your new friend SED

By : Brittany Kamai
18 February 2010
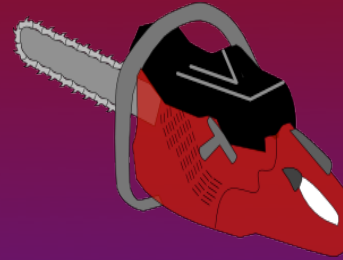
Solving a problem requires choosing the correct tool

GREP · SED · AWK

# When to use sed

- Automate editing to be performed on one or more lines

- Simplify the task of performing the same edits on multiple files

- Write conversion programs

# SED pep talk

*HURDLES* you will face

1. How to apply RegExp

2. How to use SED syntax

3. How to interact with the shell

4. The knack of script writing

**SED IS YOUR FRIEND**

# Regular Expressions Review

**^** : beginning of the line

**$** : end of the line

**.** : any single character

**\\** : escape character

(character)**\*** : matches arbitrary number of occurences

(character)**?** : matches 0 or 1 occurences

(character)**+** : matches 1 or more occurences

# [ brackets are special ]

[ ] : matches anything enclosed in the [ ]

[^abc] : excludes everything after the character

[[:alpha:]] = [a-zA-Z] returns all upper and lower case alphabets

[[:upper:]] = [A-Z] returns all upper case letters

[[:lower:]] = [a-z] returns all lower case letters

[[:digit:]] = [0-9] returns any thing with digits

[[:alnum:]] = [0-9a-zA-Z] any thing with numbers or letters (no symbols only)

[[:space:]] matches any string that has spaces in it

# [ brackets are special ]

^ : loses its special meaning in a bracket if not placed first

   [ab^] means nothing special


] : loses special meaning if placed in a list

   [ ] | 2 ] ---> returns ], 1, 2


- : loses its special meaning if placed last

    [892-] means nothing special

matching specific # of repetitions of a pattern

grep   "[[:digit:]]\{3\}[ -]\?[[:digit:]]\{4\}"   file

any number

set of 3

seperated by
0 or 1 dashes

any number

set of 4

what you are searching for
808 - 7631

# How SED works

**Input**

**Pattern Space**

**Output**

Reads from commands from standard input
or
from a file

- Streams through your file line by line looking for the correct pattern

- Performs sequence of editing commands in the pattern space

Outputs to the terminal unless otherwise specified

# SED syntax

% SED 'command' file.txt

Unix command

command you want to execute

file that you want to execute commands

# SED syntax

**Multiple commands**

% SED -e 'command1' -e 'command2'  file.txt

**Piping unix commands**

% {unix command}| sed -e 'command1' file.txt

**Reading commands from a file**

% SED -f  sedscript.sed

# Deleting Lines

% SED '1,10d' file.txt

Finds line 1 through 10 → Deletes lines 1 through 10

returns a file from 10 - N

```
1.  star 3034       11.76 +0.46 +0.09 3
2.  planet 3241     13.78 +0.57 +0.06 3
3.  STAR 149        16.42 +0.60 +0.25 1
4.  star 1593       11.12 +0.75 +0.30 3
5.  galaxy 3122     11.92 +0.73 +0.19 1
6.  starburst 778    8.90 +0.99 +0.22 4
```

**% sed -e '1,2d' file**

**% sed -e '3d' file**

```
3. STAR 149        16.42 +0.60 +0.25 1
4. star 1593       11.12 +0.75 +0.30 3
5. galaxy 3122     11.92 +0.73 +0.19 1
6. starburst 778    8.90 +0.99 +0.22  4
```

```
1. star 3034       11.76 +0.46 +0.09 3
2. planet 3241     13.78 +0.57 +0.06 3
4. star 1593       11.12 +0.75 +0.30 3
5. galaxy 3122     11.92 +0.73 +0.19 1
6. starburst 778    8.90 +0.99 +0.22  4
```

```
1.  star 3034        11.76 +0.46 +0.09 3
2.  planet 3241      13.78 +0.57 +0.06 3
3.  STAR 149         16.42 +0.60 +0.25 1
4.  star 1593        11.12 +0.75 +0.30 3
5.  galaxy 3122      11.92 +0.73 +0.19 1
6.  starburst 778     8.90 +0.99 +0.22  4
```

% sed -e '1,2d' -e '4,5d' -e 's/STAR/star/' file

```
3.  star 149         16.42   +0.60 +0.25  1
6.  starburst 778     8.90 +0.99 +0.22  4
```

```
readcol,'ple.stauffer.with.bi',stname,stname2,  stv, stbmv, stvmr, strmi,

usersym, cos(findgen(32.)*!PI*2./32.),sin(findgen(32.)*!PI*2./32.),/fill
----
plot,BKbmv,BKv2, yrange = [18,8],xrange = [0,2.],/ystyle,/xstyle,$
    xtitle = 'B-V', ytitle = 'V',charsize = 2.0,$
    color = black, background = white,/nodata
oplot, jmBMV,jmV, color = black, psym = 1, symsize = 0.75
device, /close
```

to

## % sed -e 'plot/,/\/nodata/d' file

```
readcol,'ple.stauffer.with.bi',stname,stname2,  stv, stbmv, stvmr, strmi,

usersym, cos(findgen(32.)*!PI*2./32.),sin(findgen(32.)*!PI*2./32.),/fill
----


oplot, jmBMV,jmV, color = black, psym = 1, symsize = 0.75
device, /close
```

# search and replace

% SED 's/searchIT/replaceIT/' file.txt

Line by line
finds
the first
occurence of
searchIT

→

Replaces
something
with replaceIT

replaceIT

file.txt

| | | | | | | |
|---|---|---|---|---|---|---|
| star | 2484 | file | 9.16 | +0.16 +0.15 2 | 0520222484 |
| star | 3034 | file | 11.76 | +0.46 +0.09 3 | 0520223034 |
| planet | 3241 | | 13.78 | +0.57 +0.06 3 | 0520223241 |
| STAR | 1491 | | 16.42 | +0.60 +0.25 1 | 0520221491 |
| star | 1593 | | 11.12 | +0.75 +0.30 3 | 0520221593 |
| galaxy | 3122 | | 11.92 | +0.73 +0.19 1 | 0520223122 |
| starburst | 778 | | 8.90 | +0.99 +0.22 4 | 0520223128 |

Print modified line

% SED   -e   -n  's/[a-z]*/light/p'  file.txt

Do not print
unless there is
a match

| | | | | | |
|---|---|---|---|---|---|
| light | 2484 | file | 9.16 | +0.16 +0.15 2 | 0520222484 |
| light | 3034 | file | 11.76 | +0.46 +0.09 3 | 0520223034 |
| light | 3241 | | 13.78 | +0.57 +0.06 3 | 0520223241 |
| light | 1593 | | 11.12 | +0.75 +0.30 3 | 0520221593 |
| light | 3122 | | 11.92 | +0.73 +0.19 1 | 0520223122 |
| light | 778 | | 8.90 | +0.99 +0.22 4 | 0520223128 |

file.txt

| | | | | | | |
|---|---|---|---|---|---|---|
| star | 2484 | file | 9.16 | +0.16 | +0.15 2 | 0520222484 |
| star | 3034 | file | 11.76 | +0.46 | +0.09 3 | 0520223034 |
| planet | 3241 | | 13.78 | +0.57 | +0.06 3 | 0520223241 |
| STAR | 1491 | | 16.42 | +0.60 | +0.25 1 | 0520221491 |
| star | 1593 | | 11.12 | +0.75 | +0.30 3 | 0520221593 |
| galaxy | 3122 | | 11.92 | +0.73 | +0.19 1 | 0520223122 |
| starburst | 778 | | 8.90 | +0.99 | +0.22 4 | 0520223128 |

**G**lobal

% SED -e 's/[a-z]*/light/g' file.txt

| | | | | | | |
|---|---|---|---|---|---|---|
| light | 2484 | light | 9.16 | +0.16 | +0.15 2 | 0520222484 |
| light | 3034 | light | 11.76 | +0.46 | +0.09 3 | 0520223034 |
| light | 3241 | | 13.78 | +0.57 | +0.06 3 | 0520223241 |
| STAR | 1491 | | 16.42 | +0.60 | +0.25 1 | 0520221491 |
| light | 1593 | | 11.12 | +0.75 | +0.30 3 | 0520221593 |
| light | 3122 | | 11.92 | +0.73 | +0.19 1 | 0520223122 |
| light | 778 | | 8.90 | +0.99 | +0.22 4 | 0520223128 |

the brown cat was chased by the brown dog
the brown cat was not chased by the brown dog

`%   sed   -e   '/not/s/brown/white/g'   file`

the white cat was not chased by the white dog

# delimiters

## / @ % # ; : _

- sed 's/\/usr\/bin/\/bin/g' file.txt

- sed 's#/usr/bin#/bin#g' file.txt

- sed 's_/usr/bin_/bin_g' file.txt

- sed 's:/usr/bin:/bin:g' file.txt

http://www.foo.com/mypage.html

% sed 's_http://www.foo.com/_http://bar.net/_' file.xt

http://www.bar.net/mypage.html

# Pattern Space

- Able to save up to 9 different patterns that SED finds
  - accessible using \1 - \9

- Need to save each pattern in \( character\)

sed 's/\([a-z]*\).*/\1/'

sed 's/\([a-z]*\) \([a-z]*\)/\2 \1/'

match lower case letters

matches characters after first match

stores first pattern

switches order of 1st and 2nd matched pattern

abcd123 → abcd

abcd efgh → efgh abcd

```
1167 989 835 1321 51553
4315 787 321 4510 99909
8953 153 153 1532 12325
7891 142 415 2532 51535
```

```
sed  -e  's:[0-9]*: & \& : g'  -e  's:..$: \\\:' file.txt
```

```
1167 & 989 & 835 & 1321 &  51553 \\
4315 & 787 &  321 & 4510 & 99909 \\
8953 & 153 & 153 & 1532 & 12325 \\
7891 &142 & 415 & 2532 & 51535 \\
```

# WEBSITES

1.http://www.linuxhowtos.org/System/sed_tutorial.htm

2.http://www.panix.com/~elflord/unix/sed.html *****

3.http://www.grymoire.com/Unix/Sed.html

4.http://www.cs.utk.edu/~vose/c-stuff/sed_tutorial.html

5.http://en.kioskea.net/faq/1753-tips-and-tricks

\g changes all of the patterns in each line

can add a number at the end of the
substitution to indicate you only want to
match that pattern

\1 <---inside substitution
/1 <---after the substitution

can use numbers in combo with global
command
/2g after the second one