Developing Your Ruby Data-Fu

Master of File Munging

Define The Problem

Do you...

Have lots of data in text files

Use lots of unix style utilities

Use Ruby?

Data Manipulation

'SQL' Like Operations in the Shell

Meet the Gang

$_

-paine

$F

CAT

ruby -pe '$_' /etc/passwd

CAT

`ruby -pe '$_' < /etc/passwd`

CAT

cat /etc/passwd | ruby -pe '$_'

CUT

`ps wwaux | ruby -ane 'puts $F[0]'`

TR

cat some-lines.txt |

ruby -pe '$_.gsub! /[0-9]/,"*"'

SELECT X,Y,Z

ruby -ane 'puts $F[0..2].join("\t")'

SELECT * WHERE ?

```
ruby -ne 'puts $_ if /ohai/i'
```

SELECT * WHERE ?

ruby -ane 'puts $_ if $F[0] =~ /nose/i'

WGET

ruby -r open-uri -e 'puts open(ARGV[0]).read'

TR

ruby -pe '$_.tr "a-z","A-Z"'

AUTOINCREMENT

```
ruby -ne 'puts "#$.\t#$_"'
```

Print Lines 5-10

ruby -ne 'puts $_ if $.<=10 && $.>=5'

Print Lines 5-10

```
ruby -ne 'puts $_ if (5..10).include? $.'
```

Edit Multiple Files at Once

```
ruby -i.bak -pe '$_.gsub! "2000-2008", "2000-2009"' *.rb
```

Things to Add

- **-0sep**

- **-F pattern**

- **-i.bak** (edit files in place w/backup)

- **-r libname** (and find common useful libs)

- sum, min, max

- appending on a computed column

Things to Add

- rec-view.rb

- record stream libraries (you still need to write them)

- processing all the files in a directory to do a similar task (eg: to change the copyright comment from 2008 to 2009? - good reason to show -i.bak)