

le système unix et statistiques

Joseph Saint Pierre

Université de Toulouse

rencontre ingénieurs statisticiens 15 octobre 2012

Analogies et histoire

Bell Labs, Essentiel dans le développement de Unix et du langage statistique **S** et donc implicitement de **Splus** et de **R**

ressemblance des commandes `rm`, `ls - rm()`, `ls()`

Le langage **S** a été très vite porté sur unix

Influence de **GNU** tant sur **Unix** que sur **R**

Commandes unix utiles

file donne le type du fichier

wc compte les lignes, mots et caractères du fichier

wc -l compte seulement les lignes du fichier

head, tail affiche le début ou la fin d'un fichier

Importance des options dans unix

head -5 affiche les 5 premières lignes

tail -2 affiche les 2 dernières lignes

tail -r affiche le fichier changeant l'ordre

Des commandes très importantes

cut Affiche des colonnes ou des champs d'un fichier

cut -f Pour les champs

cut -c pour les colonnes

grep Affiche des lignes contenant des expressions.

base de données du pauvre...

De très nombreuses options :

grep -v affiche lignes ne contenant pas

grep -n affiche le numéro des lignes

grep -l affiche les noms des fichiers contenant les expressions.

Expressions régulières

^\$.*?

sort tri des lignes d'un fichier

De nombreuses options.

D'autres commandes utiles

nl numérote les lignes

uniq pour détecter les lignes doublons, les enlever

paste pour coller des fichiers, ajout de colonnes

cat pour coller des fichiers, ajout de lignes

join pour joindre des fichiers avec une clef

fold pour replier les lignes trop longues

Quelques points sur unix

redirection

sort ii > jj envoie le fichier ii trié dans le fichier jj

sort ii >> jj envoie le fichier ii trié à la fin du fichier jj

tube de communication

Essentiel dans le fonctionnement du système.

grep text ii | sort | uniq

Les commandes travaillent simultanément.

tâches de fond

grep text ii | sort | uniq >>jj &

interactif

On peut utiliser le fichier jj en cours de commande !

Éditeurs sed awk perl

Exemples d'utilisation remplacer les ; par des espaces avec sed

```
sed 's/;/ /g' dd > dd2
```

Insérer une ligne blanche entre chaque ligne d'un fichier avec sed

```
sed G dd
```

avec awk

```
awk '{print $0 "\n"}' dd
```

perl est encore plus puissant mais plus bavard.

Aller plus loin

Aide en ligne commandes unix **man** et **info**.

Beaucoup de documentations et d'exemples sur Internet

Comprehensive Perl Archive Network,

[http ://www.cpan.org/](http://www.cpan.org/)

Questions, exemples, formations...