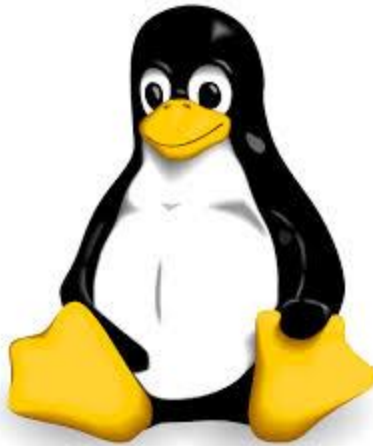


Unix Commands



Jeremiah Lant, Hydrologist
USGS Kentucky Water Science Center
jlant@usgs.gov

Last meeting

- Learned some more Unix commands
 - Creating aliases and variables; **.bashrc**
 - Find matching text in files; **grep**
 - **grep** (global/regular expression/print) can also match patterns using regular expressions
 - Find files themselves whose names match a given pattern; **find**

Today's Objective

1. **More hands-on learning** with Unix commands
 - Finish trying out commands **Cut out columns** from a data file; **cut**
 2. **Introduce the for loop;** A loop that is executed once for each value in some kind of set, list, or range.
-
- **Why?**
 - Good basis for learning how to program.
 - Becoming more comfortable on command line
 - Becoming more efficient

Answers to Meeting 6 Questions

1. How to combine patterns using **grep**?

Use the **-e** flag.

```
$ grep -v -e “#” -e 5s test.txt | head
```

2. How to combine **mv** and **find** together?

Use **\$(command)** which inserts a command's output in place.

```
$ mv $(find . -name *.txt) temp/
```

3. What is the difference between **history** command and **.bash_history** file?

Try Out Commands – page 1

make a alias for the list command that lists in long human-readable and time ordered format

\$ alias ls="ls -lht"

create an environment variable called MYDATA for data directory in meeting-5-0

\$ export MYDATA=/c/Users/jlant/meeting-5-0/data

\$ echo \$MYDATA

open another shell instance.

Does \$MYDATA exist in the other shell instance? Why or why not?

No, because environment variable not in the .bashrc file.

Using data in meeting-5-0 directory to answer the following:

What was the discharge value 1 year ago to this day in test.txt

\$ grep 2013-04-30 test.txt

Using data in meeting-5-0 directory to answer the following:

Remove header from test.txt file and create a new file called "test_without_header.txt"

\$ grep "#" -v test.txt

Try Out Commands – page 2

Is there a file called test2_uv.txt in the meeting-5-0 directory?

What is the path to the file from your home directory?

\$ /c/Users/jlant/meeting-5-0/data/test-files/test2_uv.txt

From your home directory, find how many unit value files there are in the meeting-5-0 directory.

\$ find meeting-5-0 -name *_uv.txt | wc -l

From your home directory, find how many total text files are there in meeting-5-0.

\$ find meeting-5-0 -name *.txt | wc -l

Try Out Commands – page 3

How many instances does the number 25 occur in the data file called discharge-week1.txt in the meeting-5-0/data/sample-weekly-discharge directory?

How many instances does the number 25 occur in all the data files in the meeting-5-0/data/sample-weekly-discharge directory?

What is the smallest discharge in the data file called discharge-week1.txt in the meeting-5-0/data/sample-weekly-discharge directory?

What are the 2 largest discharges in all the data files in the meeting-5-0/data/sample-weekly-discharge directory?

What is the smallest discharge in all the data files in the meeting-5-0/data/sample-weekly-discharge directory?

Try Out Commands – page 4

make a directory called meeting-7/data; **mkdir -p meeting-7/data**
make a file called stage.txt in meeting-7/data that has the following format:

Date, Parameter, Value
2014-04-29, stage, 17
2014-04-30, stage, 20

\$ cut -d , -f 1 stage.txt **# cuts by comma delimiter and gets column (field -f) 1**

\$ cut -d , -f 1-3 stage.txt **# cuts by comma delimiter and gets columns (field -f) 1-3**

make a tab-delimited file of stage.txt; **tab-delimited is default**

\$ cut -f 1 stage.txt

\$ cut -f 1-3 stage.txt

Try Out Commands – page 5

copy 03290500_dv.txt from meeting-5-0/data/nwis_files/ to meeting-7-0

find the largest discharge value in April of 2013

find the largest discharge value in entire file

```
$ grep -v "#" 03290500_dv.txt | grep -v agency_cd | grep -v 5s | cut -f 4 | sort -n  
| uniq | tail -1
```

OR

find the largest discharge value along with the date in entire file

OR

Try Out Commands – page 7

what if you wanted to know the largest value for multiple daily value files?

for loop - A loop that is executed once for each value in some kind of set, list, or range.

create 5 dummy files called data1.txt, data2.txt, etc.

rename a lot of files

for filename in *.txt

do

mv \$filename processed-\$filename

done

Try Out Commands – page 8

what if you wanted to know the largest value for multiple daily value files?

for loop - A loop that is executed once for each value in some kind of set, list, or range.



Next meeting

- Make a bash script
- Learning version control with Git