

Lecture 11 (2-6-26)

Pipelines

Motivating Questions

1. Why is the pipeline pattern so powerful?
 2. How do we use regular expressions to match text?
 3. How do we translate characters?
 4. How do we extract fields?
 5. How do we search for patterns?
 6. How do we modify text streams?
 7. How do we perform more complex text processing?
-

- `cut` and `awk` to select **particular fields** in the text
-

investigating the CSE sample curriculum via `curl`

- How can I extract lines that have "Math", "Physics", or "CSE"
 - `grep -E (MATH|PHYS|CSE)`
- But this also gave me the twitter handle
 - So I add a `grep`
 - `grep -v meta | grep credits | grep -Eo '(MATH|PHYS|CSE|) | sort | uniq -c`
 - `grep -v meta` ignores the meta tag data
 - `grep credits` so i get rid of the twitter handle (since all courses are followed by "x credits")
- getting total credit hours
 - ``grep -v meta | grep -E 'Total Credit Hours: [0-9.]+'`
- to get the just the numbers you can **add any of the following:**
 - `grep -Eo '[0-9\.]+'`
 - `awk '{print $4}'`
 - `cut -d ' ' -f 4`

- or you could use `sed` to do the whole thing
 - `sed -En 's/.*Total Credit Hours: ([0-9\.]+).*/\1/p'`
-

Infamous Old Test Question

```
#!/bin/sh  
  
echo "Hello $USER"
```

- this doesn't echo "Hello \$USER"
 - you can use `cat -Ev myShellScript.sh`
 - to see all the hidden characters and the end of lines
 - there was some windows `'r'` issue with this
 - to fix you can
 - `cat ./hello.sh | tr -d '\r' > hello.sh.fixed`
 - then do `chmod +x hello.sh.fixed`
 - then run `./hello.sh.fixed` and it will run
-

- remember to call a function in shell you just type the name in the script
 - also remember that if you want to consume the argument a flag then you need to double shift -> used this on homework03
-

replace space in "New York" with "%20"

- `sed 's/ /%20/g'`