

Unix Tools
Courant Institute of Mathematical Sciences
Homework assignment 2
Due: March 11, 2007

1. Download the data using the command `curl`:

```
$ curl "http://www.usatoday.com/news/politicselections/vote2004/\
PresidentialByCounty.aspx?oi=P&rti=G&tf=1&sp=FL" > \
fl.county.2004.html
```

How would the shell interpret this if you did not use quotes around the URL?

2. Use a pipe of `awk` and `sed` with no more than one instance of each to extract county level data from the HTML file and write it to file `fl.county.2004`. The result should have six columns separated with the separator `*`. Why is space not a good separator here?
3. Use `awk` to determine how many precincts are there in FL and how many reporting.
4. Use `awk` to find how many votes were registered for Bush, Kerry, and Nader.
5. Use `awk` to find out how many counties did Bush win.
6. Use `awk` and `sort` to determine who won the largest county.
7. Use `awk` and `sort` to determine how far down in the list of counties sorted in decreasing order of total number of votes we need to go to find a county Bush won.
8. Download the Florida 2000 Presidential elections data set from

```
http://www.stat.unc.edu/faculty/rs/source/Data/fl.dat1.txt
```

The row starting with 50 gives the results for the Palm Beach County.

9. Use `awk` to print the count and the county number of the county where Buchanan obtained the largest number of votes.

10. Use **awk** to determine the average number of votes Buchanan obtained in Florida.
11. Use **awk** to determine the standard deviation σ of Buchanan's votes in Florida, that is the square-root of the average of $(|\text{votes}| - |\text{average_votes}|)^2$.
12. Use **awk** to find out the number of standard deviations separating the number of votes Buchanan obtained in Palm Beach from the average number of votes he obtained in Florida, in absolute value.
13. Use **awk** and **sort** to print that number in decreasing order for all counties.
14. Use **awk** determine by how many votes V Bush won in Florida. Give the ratio V/σ .