GENETICS

NHANES R DATA PARSER

12.02.2007

NHANES is a great ressource for doing epidemiological research. As the NIH website provides only data import for commercial software here is my rewrite in R. First load from their site

adult.exe
youth.exe
lab.exe
lab2.exe
exam.exe

put everything in one directory and expand the self-extracting archives. Then create from each SAS file a new variable content file that will only contain variable name and tab separated start position in the .dat file. Adult.var for example would read like this:

```
SEQN 1
DMPFSEQ 6
DMPSTAT 11
DMARETHN 12
DMARACER 13
...
HAZNOK5R 3345
```

Then start the following R job with the datasets and variables that you are interested in

|wj_nhanes.R|