Drawing a Random Sample with SPSS[1]

Sometimes it is necessary or useful to select a random sample from your data. Sometimes a specific number of cases is required, and sometimes rough percent is needed. The following SPSS programs will show how to select either type.

```
* random_sample.sps .
** Create 100000 cases for this example .
INPUT PROGRAM.
LOOP #I=1 TO 100000.
COMPUTE case = 1 .
END CASE.
END LOOP.
END FILE.
END INPUT PROGRAM.
EXECUTE.

* Create a random variable .
* The following command will create a random variable with .
* values ranging from 1 to 100.
** If you want to be able to replicate your sample - set the "seed" .
** to some nomber .
set seed 10 .
compute randvar = uniform(100) .
save outfile = 'c:\tr\totalN.sav' .

** Demonstrate how to get an approximate 20% sample .
get file =  'c:\tr\totalN.sav' .
** If I don't need an exact N I can select based on this variable .
select if randvar ge 0 and randvar le 20 .
desc var = all .


** Demonstrate how to get an EXACT 20% sample (20,000 of 100,000 cases .
get file =  'c:\tr\totalN.sav' .
** If I don't need EXACT N I can sort by randvar,
** number the cases, and the select the first case .
** Since the cases will be in random order, it will be a random selection .
sort cases by randvar .
compute idvar = 1 .
* Use "create" and "csum" to create an incremented variable.
* The first case will be 1, second 2, and so on.
create idvar = csum(idvar) .
variable label idvar 'Identification Variable' .
select if idvar le 20000 .
desc var = all .
```

[1] Prepared by Patty Glynn, University of Washington, March 15, 2006. C:\all\help\helpnew\RandomSampleSPSS.wpd.