

```
*****  
*****
```

```
**** Poiminta - 2. vaihe
```

```
*****  
*****
```

```
*****  
*****
```

```
**** Asetetaan satunnaislukugeneraattorin siemenluku koneajasta
```

```
*****
```

```
global system_time = clock("`c(current_date)`c(current_time)", "DMYhms")
```

```
global seed=substr("$system_time",1,10)
```

```
set seed $seed
```

```
**** 1. Koeryhmän poiminta
```

```
*****  
*****
```

```
** Ladataan aineisto
```

```
use "${INFOLDER}\full_targetgroup.dta", clear
```

```
** Stratifioitu otanta
```

```
count
```

```
global n1p=100*(31000/`r(N)')
```

```
sample ${n1p}, by(stratacell)
```

```
save "${INFOLDER}\treatment_group_initial.dta", replace
```

```
global NT_initial = _N
```

** Stratifioitu otanta voi antaa pyöristyvirheestä johtuen hieman ryhmän tavoitekoosta poikkeavan havaintomäärän.

** Täsmäys tehdään tarvittaessa suoralla otannalla.

** Jos stratifioitu otanta antoi täsmälleen 31000 havaintoa, tallennetaan taulu

```
if $NT_initial == 31000 {
```

```
    gen treatment=1
```

```
    save "${OUTFOLDER}\treatment_group.dta", replace
```

```
}
```

** Jos stratifioitu otanta antoi yli 31000 havaintoa, poistetaan havainnot suoralla otannalla

```
if $NT_initial > 31000 & $NT_initial != . {
```

```
    sample 31000, count
```

```
    gen treatment=1
```

```
    save "${OUTFOLDER}\treatment_group.dta", replace
```

```
}
```

** Jos stratifioitu otanta antoi alle 31000 havaintoa, täydennetään havainnot suoralla otannalla jäljelle jääneistä yrityksistä

```
if $NT_initial < 31000 {
```

```
    use "${INFOLDER}\full_targetgroup.dta", clear
```

```
    merge 1:1 ${id} using "${INFOLDER}\treatment_group_initial.dta", gen(match) keep(match  
master)
```

```

drop if match==3
drop match

global Nadd1 = 31000 - $NT_initial
sample $Nadd1 , count

append using "${INFOLDER}\treatment_group_initial.dta"

gen treatment=1
save "${OUTFOLDER}\treatment_group.dta", replace
}

* Poistetaan väliaikainen tiedosto

erase "${INFOLDER}\treatment_group_initial.dta"

**** 2. Kontrolliryhmän muodostavat jäljelle jäävät yritykset
*****

** Ladataan aineisto

use "${INFOLDER}\full_targetgroup.dta", clear

** Poistetaan koeryhmään poimitut ja tallenetaan taulu

merge 1:1 ${id} using "${OUTFOLDER}\treatment_group.dta", nogen keep(match master)
keepusing(treatment)

drop if treatment==1

drop treatment

```

```
gen treatment=0
```

```
save "${OUTFOLDER}\control_group.dta", replace
```