

Sociology 213 Lab4

Lab 4: In this lab we will aggregate several years of state-level census data together, make a map of % black by state in 1880, 1910, and 1940, and do a descriptive analysis of IPUMS data on blacks in 1910 and 1940.

Assignment for Lab4 (due next week Friday)

Note: If you are having difficulty, ask for my help—that is the nice thing about a small class. This assignment should be fun—as we are doing hands on social research—not frustrating.

1. Add new data from 1880, 1910, and 1940 to Arcview: Run aggstate.do and add the resulting data to the map states.shp (the new state map) using dbmscopy to convert it to dbf format. Make maps of the percent black by state in 1880, 1910, and 1940.
2. Analyze data from the 1910 and 1940 IPUMS data: Look at my results in the do-file lab4.do. Using the resulting data set (bl1040.dta) make a do-file that provides two interesting pieces of information relating to the migration, geographic distribution and/or socioeconomic situation of blacks in 1910 to 1940. I.e., you have done some background reading on the Great Migration of blacks from the south to the north. Use the IPUMS data to add some demographic details to that story.

Give me a hardcopy of the do-file and the interesting part of the results (the log or smcl file)—i.e. edit the log file. Write a few sentences explaining why the results are interesting.

Example of a do file:

-----mydofile.do-----

```
use bl1040
des
sum
tab x y
tab x y if g==94
tab x y, row nofreq
tab occat region2 if year==94, row nofreq
```

etc.

```

-----Lab4.do-----

clear
use 1940nw
append using ipums1910

keep if raceg==2
tab year

gen samest=bplg==state
tab samest year
tab samest year [w=perwt], col nofreq

gen illinois=state==17
tab ill year
gen m_chi=metaread==1600
gen m_det=metaread==2160
gen m_ny=metaread==5600

tab year m_chi
tab year m_det
tab year m_ny

gen bigcity=metaread>0 & metaread~=.
tab year bigcity, row nofreq

* notice how many people are living in the same state they were born

gen age2=autocode(age,5,0,100)
tab age2

gen occat=occl1950

recode occat 0/99=1 100/123=2 200/290=3 300/390=4 400/490=5 500/595=6
/*
*/ 600/690=7 700/790=8 800/840=9 900/970=10 980/998=11 999=.

* note: 999=not working (blank)

lab define occat 1 "professional and technical" 3 "managers" 4
"clerical" /*
*/ 2 "agriculture" 6 "craftsman" 5 "sales" 7 "operatives" 8 "service
workers" /*
*/ 9 "farm laborers" 10 "laborers" 11 "misc"

lab val occat occat

gen region2=region
recode region2 11/13=1 21/23=2 31/33=3 41/43=4 91/99=9
lab define region2 1 "northeast" 2 "midwest" 3 "south" 4 "west" 9
"unknown"

```

```

lab val region2 region2

sort year
by year: tab region2 bigcity, row nofreq

* we want to make a region2 variable for the place of birth

save temp, replace

collapse region2, by(state)
rename state bplg
rename region2 bpreion
sort bplg
save bpcodes, replace

use temp
sort bplg
merge bplg using bpcodes
tab _merge
replace bpreion=. if _merge==1
label val bpreion region2
drop if _merge==2

tab bpreion

sort year
tab region2 bpreion
by year: tab region2 bpreion [w=perwt], row nofreq

tab region2 bpreion if age2==20 & year==94 [w=perwt], row nofreq
tab region2 bpreion if age2==40 & year==94 [w=perwt], row nofreq
tab region2 bpreion if age2==80 & year==94 [w=perwt], row nofreq

tab occat
tab occat year [w=perwt], col nofreq
sort year
by year: tab occat region2 [w=perwt], col nofreq

by year: tab occat region2, col nofreq
* note that the 1910 file should have perwt

save bl1040, replace

```

```

-----aggstate.do-----
* note: all three of these files have 48 cases
capture log close
log using aggstate, replace
set more off
clear

use s1880map
rename total p1880
rename white w1880
rename colored b1880
rename pctblk pb1880
rename state st1880
keep id p1880 w1880 b1880 pb1880 st1880
sort id
save d, replace

clear

use state2
gen str3 sta=substr(state,-3,3)
gen str3 stb=substr(state,1,3)
gen str6 st2=sta+stb

sort st2
save state3, replace

use s1910

rename state st1910
rename topo p1910
rename wpop w1910
rename bpop b1910
rename pctblk pb1910
drop pctwt
gen n=_n
sort n
des
save a, replace

use s1930
rename total p1930
gen w1930=whitemen+whitewom
gen b1930=blkmen+blkwom
gen pb1930=b1930/p1930
rename state st1930
keep p1930 w1930 b1930 p* st
gen n=_n
sort n
des

```

```
save b, replace

use s1940
rename total p1940
gen w1940=tot_wht
gen b1940=tot_blk
gen pb1940=b1940/p1940
rename state st1940
keep p1940 w1940 b1940 p* st
gen n=_n
sort n
des
save g, replace
```

```
use s1950
gen p1950=totmale+totfem
gen w1950=nwm+fwm+nwfm+fwfm
gen b1950=negmale+negfem
gen pb1950=b1950/p1950
rename state st1950
keep p1950 w1950 b1950 p* st
gen n=_n
sort n
save c, replace
```

```
use a
merge n using b
tab _merge
drop _merge
sort n
merge n using c
drop _merge
sort n
merge n using g
tab _merge
drop _merge
```

```
* check the match
list st*
```

```
* type lookup abbrev, then help functions
```

```
gen str3 sta=substr(st1910,-3,3)
gen str3 stb=substr(st1910,1,3)
gen str6 st2=sta+stb
```

```
replace st2=lower(st2)
```

```
list st2
sort st2
capture drop _merge
```

```
merge st2 using state3

tab _merge

list st2 state st1910

drop _merge

sort id

merge id using d

tab _merge

sum

* check to see if total population matches

egen up1910=sum(p1910)
egen up1930=sum(p1930)
egen up1880=sum(p1880)
egen up1940=sum(p1940)
egen up1950=sum(p1950)

egen bp1910=sum(b1910)
egen bp1930=sum(b1930)
egen bp1940=sum(b1940)
egen bp1880=sum(b1880)
egen bp1950=sum(b1950)

for var up1910-bp1950: replace X=X/1000000

list up* bp* in 1

sort id

save aggstate, old replace
* note: save in old format so that dbms copy can read it

! del a.dta
! del b.dta
! del c.dta
! del d.dta
! del g.dta
```

-----ipums.do-----

```
clear
capture log close
set more 1
set mem 30m
log using ipums, replace
```

```
use ipums1910
gen occ2=occ
```

```
do occ1910
```

```
tab raceg
tab raceg if occ>0
```

```
gen white=raceg==1
gen black=raceg==2
```

```
gen male=sex==1
gen female=sex==2
```

```
* table occ, c(sum male sum female)
```

```
* table occ, c(sum white sum black)
```

```
table occ if occ<100, c(sum white sum black)
```

```
collapse occ2 (sum) white black, by(occ)
```

```
gen n=white+black
gen pctblk=black/(white+black)
* egen totblk=sum(black)
* gen pctofb=black/totblk
sort pctblk
```

```
format occ %20.0g
```

```
list occ occ2 n black pctblk
```