Lab 9: Social Network Data

So far in Project #2 we have studied the demographics of residential segregation. In this lab, we are going to investigate social segregation—i.e., the demographics of friendship segregation, by race and by gender. Social network data poses several problems in data management. There are two types of social network data:

1) Characteristics of the friends are part of each respondent's case.
Example: the MCSUI data
Data on up to 3 friends for each respondent. Sex and race variables for each friend (as well as other variables) are separate variables.

variable name	storage type	display format	value label	variable label
hsexper1	int	%8.0g	hsexper1	person 1: sex
hsexper2	int	%8.0g	hsexper2	person 2: sex
hsexper3	int	%8.0g	hsexper3	person 3: sex
hracper1	int	%8.0g	hracper1	person 1: race or ethnicity
hracper2	int	%8.0g	hracper2	person 2: race or ethnicity
hracper3	int	%8.0g	hracper3	person 3: race or ethnicity

2) Each respondent's case gives the ID's of the friends.
Example: the Add Health data
Contains data from sfriend2.dta
obs: 90,118
vars: 11
size: 8,290,856 (73.6% of memory free)

variable name	storage type	display format	value label	variable label
aid	str8	%8s		RESPONDENT IDENTIFIER
mf1aid	double	%10.0g		MALE FRIEND 1
mf2aid	double	%10.0g		MALE FRIEND 2
mf3aid	double	%10.0g		MALE FRIEND 3
mf4aid	double	%10.0g		MALE FRIEND 4
mf5aid	double	%10.0g		MALE FRIEND 5
ff1aid	double	%10.0g		FEMALE FRIEND 1
ff2aid	double	%10.0g		FEMALE FRIEND 2
ff3aid	double	%10.0g		FEMALE FRIEND 3
ff4aid	double	%10.0g		FEMALE FRIEND 4
ff5aid	double	%10.0g		FEMALE FRIEND 5

In this lab we will use the MCSUI data. We will transform the individual level data into a data set of friendship dyads, that has information on the characteristics of the respondents and their friends. To apply these techniques to network data such as Add Health, you would need to go one step further than what we will do here: you would need to merge the dyad data to the original data to get the characteristics of the friends.

We will create our dyad data using the reshape command in Stata.

Reshape converts data from wide to long form and vice versa. Think of the data as a collection of observations x_ij. One such collection might be

(wide form)

-i-		------- x_ij --------		
id	sex	inc80	inc81	inc82
1	0	5000	5500	6000
2	1	2000	2200	3300
3	0	3000	2000	1000

(long form)

-i-	-j-	-x_ij-	
id	year	sex	inc
1	80	0	5000
1	81	0	5500
1	82	0	6000
2	80	1	2000
2	81	1	2200
2	82	1	3300
3	80	0	3000
3	81	0	2000
3	82	0	1000

This describes what we will do conceptually.

We will go over the file lab9.do in detail.

Lab 9 overview

1. Run lab9.do, generating the social network data—in dyad form--(friends.dta) and recoding the original MCSUI dta (saved as mcsui2.dta)

2. Learn how to make tables from Stata output (leftover from lab 8)
2a. Import a log file into Excel.
2b. Use outreg with regression commands to import formatted tables.

3. Work on your own on the following:
a. How much gender segregation of friendships is there? Make a table showing the percentage of men's friends that are male, and the percentage of women's friends that are female.
b. How much does the level of racial friendship segregation differ in LA, Boston, and Atlanta? Make a table for each of the three cities. How does this relate to the different levels of residential segregation in these three cities?

Do one of the following (i.e. either c1, c2, or c3):
c1. With the aggregate measures of the number of friends by race (f_w, f_b, f_h, and f_a) see if you can answer this question: Do whites who have friends that are minorities exhibit more tolerant racial attitudes? Choose one or two attitude questions and see if this is true.

c2. Do blacks and Hispanics who have white friends have higher wages? Are they more likely to use a white friend to find work?

C3. Is there a relationship between having interracial friends and reports of race/ethnic discrimination at work? [see question F32, fwkdiscr]

Homework:
Do Part 3, a, b, and one of c1, c2, c3. Make at least one formatted table based on the technique we learned in lab.

```
------------------lab9.do---------------------------------
set more off
clear
capture log close
log using lab9, replace text


use mcsui_hh2


* gender
rename crespsex sex
lab define sex 1 "male" 2 "female"
lab val sex sex
tab sex

* label the race variable
gen race=gicrace
recode race 1=2 2=3 3=4 4=1
lab define race 1 "white" 2 "black" 3 "hispanic" 4 "asian" 5 "other" 9
"missing"
lab val race race

tab race
drop if race>4

* label the city variable
lab define city 1 "detroit" 2 "atlanta" 3 "los angeles" 4 "boston"
lab val city city

tab city

* education
gen edyrs=eeduyrs
recode edyrs 95/100=.
sum edyrs
tab race, sum(edyrs)


* first do wages

tab wageflag

replace fhrwage=. if wageflag>0

sum fhrwage
* fhrwage is hourly wages

tab race, sum(fhrwage)

sort caseid
save temp, replace

keep hsexper* hrelper* hmarper* hneiper* hhlpper* hcntper* caseid /*
*/ hjobper* hwelper* heduper* hracper* race sex city hnumnetw
```

```
des



reshape long hsexper hrelper hmarper hneiper hhlpper hcntper /*
*/ hjobper hwelper heduper hracper, i(caseid) j(friend)

des

drop if city==1
* social network questions were not asked in Detroit

* drop if crace==5
recode hracper 7=5 8=5 9=5

drop if hrelper==0
* this commands gets rid of missing cases--"logical skips"

* drop if hrelper==1

recode hjob 7=5 8=5 9=.
recode heduper 7=. 8=. 9=.
recode hmarper 8=. 9=.
mvdecode hrelper hneiper hhlpper hcntper hwelper, mv(6)
mvdecode hrelper hneiper hhlpper hcntper hwelper, mv(8)
mvdecode hrelper hneiper hhlpper hcntper hwelper, mv(9)


lab val hracper race


lab define hrelper 1 "relative" 2 "friend" 3 "co-worker" 4 "other"
recode hrelper 5/9=4
lab val hrelper hrelper

tab friend

tab hrelper

tab race hracper if hrelper>1, row nofreq

sum heduper

gen f_w=hracper==1
gen f_b=hracper==2
gen f_h=hracper==3
gen f_a=hracper==4

save friends, replace

collapse (mean) f_ed=heduper (sum) f_*, by(caseid)

sort caseid

merge caseid using temp
tab _merge
```

```
for var f_ed f_w f_b f_h f_a: replace X=. if _merge==2


* job contacts

recode ffindhow 7/9=. 0=.
lab define ffindhow 1 "friends or relatives" 2 "other persons" 3 "newspaper" 5
"other"
lab val ffindhow ffindhow

recode fwhohelp 7/9=. 0=.
lab define fwhohelp 1 "relative" 2 "friend" 3 "acquiantance" 4 "someone else"
lab val fwhohelp fwhohelp

recode fhlprace 5/9=. 0=.
lab val fhlprace race

recode fhlpsex 0=. 7/9=.
lab val fhlpsex sex

tab ffindhow

tab race fhlprace, row

tab sex fhlpsex, row


gen wgt=wpstpere


save mcsui2, replace



keep if city>1
sum race sex edyrs f_*
egen r=rmiss(race sex edyrs f_*)
drop if r>0


xi: regress fhrwage i.race [w=wgt]
outreg using mod1, se replace

xi: regress fhrwage i.race i.sex [w=wgt]
outreg using mod1, se append

xi: regress fhrwage i.race i.sex edyrs [w=wgt]
outreg using mod1, se append

xi: regress fhrwage i.race i.sex edyrs f_ed [w=wgt]
outreg using mod1, se append

set more on
```

```
--------selected output from lab9.log-----------------------
.
. keep hsexper* hrelper* hmarper* hneiper* hhlpper* hcntper* caseid /*
> */ hjobper* hwelper* heduper* hracper* race sex city hnumnetw

.
. des

Contains data from temp.dta
  obs:         8,887
 vars:            35                          5 Mar 2002 10:33
 size:       693,186 (97.7% of memory free)
-------------------------------------------------------------------------------
-
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
-
caseid          long   %9.0g                  case number, multicity file
city            int    %11.0g      city       city
sex             int    %8.0g       sex        respondent's gender
hnumnetw        int    %8.0g       hnumnetw   number in social network
hsexper1        int    %8.0g       hsexper1   person 1: sex
hrelper1        int    %8.0g       hrelper1   person 1: how related
hmarper1        int    %8.0g       hmarper1   person 1: married
hneiper1        int    %8.0g       hneiper1   person 1: live in neighborhood
hhlpper1        int    %8.0g       hhlpper1   person 1: resp. helped
hcntper1        int    %8.0g       hcntper1   person 1: can resp. count on
hjobper1        int    %8.0g       hjobper1   person 1: has steady job
hwelper1        int    %8.0g       hwelper1   person 1: receives welfare
heduper1        int    %8.0g       heduper1   person 1: level of education
hracper1        int    %8.0g       hracper1   person 1: race or ethnicity
hsexper2        int    %8.0g       hsexper2   person 2: sex
hrelper2        int    %8.0g       hrelper2   person 2: how related
hmarper2        int    %8.0g       hmarper2   person 2: married
hneiper2        int    %8.0g       hneiper2   person 2: lives in neighborhood
hhlpper2        int    %8.0g       hhlpper2   person 2: resp. helped
hcntper2        int    %8.0g       hcntper2   person 2: can resp. count on
hjobper2        int    %8.0g       hjobper2   person 2: has steady job
hwelper2        int    %8.0g       hwelper2   person 2: receives welfare
heduper2        int    %8.0g       heduper2   person 2: level of education
hracper2        int    %8.0g       hracper2   person 2: race or ethnicity
hsexper3        int    %8.0g       hsexper3   person 3: sex
hrelper3        int    %8.0g       hrelper3   person 3: how related
hmarper3        int    %8.0g       hmarper3   person 3: married
hneiper3        int    %8.0g       hneiper3   person 3: lives in neighborhood
hhlpper3        int    %8.0g       hhlpper3   person 3: resp. helped
hcntper3        int    %8.0g       hcntper3   person 3: can resp. count on
hjobper3        int    %8.0g       hjobper3   person 3: has steady job
hwelper3        int    %8.0g       hwelper3   person 3: receives welfare
heduper3        int    %8.0g       heduper3   person 3: level of education
hracper3        int    %8.0g       hracper3   person 3: race or ethnicity
race            float  %9.0g       race
-------------------------------------------------------------------------------
-
Sorted by:  caseid
```

```
     Note:  dataset has changed since last saved

.
.
.
. reshape long hsexper hrelper hmarper hneiper hhlpper hcntper /*
> */ hjobper hwelper heduper hracper, i(caseid) j(friend)
(note: j = 1 2 3)

Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                     8887   ->   26661
Number of variables                  35   ->      16
j variable (3 values)                     ->   friend
xij variables:
             hsexper1 hsexper2 hsexper3   ->   hsexper
             hrelper1 hrelper2 hrelper3   ->   hrelper
             hmarper1 hmarper2 hmarper3   ->   hmarper
             hneiper1 hneiper2 hneiper3   ->   hneiper
             hhlpper1 hhlpper2 hhlpper3   ->   hhlpper
             hcntper1 hcntper2 hcntper3   ->   hcntper
             hjobper1 hjobper2 hjobper3   ->   hjobper
             hwelper1 hwelper2 hwelper3   ->   hwelper
             heduper1 heduper2 heduper3   ->   heduper
             hracper1 hracper2 hracper3   ->   hracper
-----------------------------------------------------------------------------

.
. des

Contains data
  obs:        26,661
 vars:            16
 size:     1,039,779 (96.5% of memory free)
-------------------------------------------------------------------------------
-
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
-
caseid          long   %9.0g                  case number, multicity file
friend          byte   %9.0g
city            int    %11.0g      city       city
sex             int    %8.0g       sex        respondent's gender
hnumnetw        int    %8.0g       hnumnetw   number in social network
hsexper         int    %8.0g       hsexper3
hrelper         int    %8.0g       hrelper3
hmarper         int    %8.0g       hmarper3
hneiper         int    %8.0g       hneiper3
hhlpper         int    %8.0g       hhlpper3
hcntper         int    %8.0g       hcntper3
hjobper         int    %8.0g       hjobper3
hwelper         int    %8.0g       hwelper3
heduper         int    %8.0g       heduper3
hracper         int    %8.0g       hracper3
race            float  %9.0g       race
```

```
-------------------------------------------------------------------------------
-
Sorted by:  caseid  friend
     Note:  dataset has changed since last saved
```

```
. tab friend

     friend |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |       4852       41.71       41.71
          2 |       3889       33.43       75.13
          3 |       2893       24.87      100.00
------------+-----------------------------------
      Total |      11634      100.00
```

```
.
. tab hrelper

    hrelper |      Freq.     Percent        Cum.
------------+-----------------------------------
   relative |       3614       31.45       31.45
     friend |       6187       53.85       85.30
  co-worker |        694        6.04       91.34
      other |        995        8.66      100.00
------------+-----------------------------------
      Total |      11490      100.00
```

```
.
. tab race hracper if hrelper>1, row nofreq

           |                        hracper
      race |     white      black   hispanic      asian      other |     Total
-----------+-------------------------------------------------------+----------
     white |     82.05       5.51       6.45       3.16       2.83 |    100.00
     black |      8.03      85.45       3.04       1.16       2.32 |    100.00
  hispanic |     13.18       4.58      78.90       1.36       1.98 |    100.00
     asian |     14.29       1.44       3.72      80.07       0.48 |    100.00
-----------+-------------------------------------------------------+----------
     Total |     38.04      29.84      19.70      10.16       2.26 |    100.00
```

```
. tab ffindhow

  how found last job |      Freq.     Percent        Cum.
---------------------+-----------------------------------
friends or relatives |       1717       49.88       49.88
       other persons |        213        6.19       56.07
           newspaper |        658       19.12       75.19
               other |        854       24.81      100.00
---------------------+-----------------------------------
               Total |       3442      100.00
```

```
.
. tab race fhlprace, row

           |          who helped get job: race
      race |     white      black   hispanic      asian |     Total
-----------+--------------------------------------------+----------
     white |       379         27         18         14 |       438
           |     86.53       6.16       4.11       3.20 |    100.00
-----------+--------------------------------------------+----------
     black |        58        490         15          7 |       570
           |     10.18      85.96       2.63       1.23 |    100.00
-----------+--------------------------------------------+----------
  hispanic |        69         17        586          4 |       676
           |     10.21       2.51      86.69       0.59 |    100.00
-----------+--------------------------------------------+----------
     asian |        22          0          6        187 |       215
           |     10.23       0.00       2.79      86.98 |    100.00
-----------+--------------------------------------------+----------
     Total |       528        534        625        212 |      1899
           |     27.80      28.12      32.91      11.16 |    100.00


.
. tab sex fhlpsex, row

           |  who helped get job:
respondent |          sex
 's gender |      male     female |     Total
-----------+----------------------+----------
      male |       724        150 |       874
           |     82.84      17.16 |    100.00
-----------+----------------------+----------
    female |       323        714 |      1037
           |     31.15      68.85 |    100.00
-----------+----------------------+----------
     Total |      1047        864 |      1911
           |     54.79      45.21 |    100.00

. keep if city>1
(1543 observations deleted)

. sum race sex edyrs f_*

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        race |    7344    2.235975   1.029859          1          4
         sex |    7344    1.595044   .5062135          1          9
       edyrs |    7321    12.20844   3.520547          0         17
        f_ed |    4635    3.970766   1.513052          1          6
         f_w |    4852    .8645919   1.168899          0          3
         f_b |    4852    .7893652   1.116649          0          3
         f_h |    4852    .4604287   .9319676          0          3
         f_a |    4852    .2390767   .7208459          0          3

. egen r=rmiss(race sex edyrs f_*)
```

```
. drop if r>0
(2715 observations deleted)

.
.
. xi: regress fhrwage i.race [w=wgt]
i.race            _Irace_1-4          (naturally coded; _Irace_1 omitted)
(sum of wgt is   5.7244e+06)

      Source |       SS       df       MS              Number of obs =    3189
-------------+------------------------------           F(  3,  3185) =   56.12
       Model |   20802.043     3  6934.01435           Prob > F      =  0.0000
    Residual |  393501.578  3185  123.548376           R-squared     =  0.0502
-------------+------------------------------           Adj R-squared =  0.0493
       Total |  414303.621  3188  129.957221           Root MSE      =  11.115

------------------------------------------------------------------------------
     fhrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Irace_2 |  -3.795313   .6121047    -6.20   0.000    -4.995472   -2.595154
    _Irace_3 |  -6.323017   .5186874   -12.19   0.000    -7.340012   -5.306022
    _Irace_4 |   .4851106   1.155871     0.42   0.675    -1.781216    2.751437
       _cons |   15.89909   .2416894    65.78   0.000     15.42521    16.37298
------------------------------------------------------------------------------

. outreg using mod1, se replace

.
. xi: regress fhrwage i.race i.sex [w=wgt]
i.race            _Irace_1-4          (naturally coded; _Irace_1 omitted)
i.sex             _Isex_1-9           (naturally coded; _Isex_1 omitted)
(sum of wgt is   5.7244e+06)

      Source |       SS       df       MS              Number of obs =    3189
-------------+------------------------------           F(  5,  3183) =   56.64
       Model |  33849.0335     5  6769.80671           Prob > F      =  0.0000
    Residual |  380454.587  3183  119.527046           R-squared     =  0.0817
-------------+------------------------------           Adj R-squared =  0.0803
       Total |  414303.621  3188  129.957221           Root MSE      =  10.933

------------------------------------------------------------------------------
     fhrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Irace_2 |   -3.58999    .602392    -5.96   0.000    -4.771105   -2.408874
    _Irace_3 |  -6.463337   .5103651   -12.66   0.000    -7.464014   -5.462659
    _Irace_4 |   .5402715   1.136922     0.48   0.635    -1.688903    2.769446
     _Isex_2 |  -4.050073   .3876519   -10.45   0.000    -4.810146       -3.29
     _Isex_9 |  -2.414163   15.99308    -0.15   0.880    -33.77194    28.94361
       _cons |    17.9314   .3071819    58.37   0.000     17.32911     18.5337
------------------------------------------------------------------------------

. outreg using mod1, se append
```

```
.
. xi: regress fhrwage i.race i.sex edyrs [w=wgt]
i.race            _Irace_1-4          (naturally coded; _Irace_1 omitted)
i.sex             _Isex_1-9           (naturally coded; _Isex_1 omitted)
(sum of wgt is   5.7244e+06)

      Source |       SS       df       MS              Number of obs =    3189
-------------+------------------------------           F(  6,  3182) =   78.50
       Model |  53418.9172     6  8903.15286           Prob > F      =  0.0000
    Residual |  360884.703  3182  113.414426           R-squared     =  0.1289
-------------+------------------------------           Adj R-squared =  0.1273
       Total |  414303.621  3188  129.957221           Root MSE      =   10.65

------------------------------------------------------------------------------
     fhrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Irace_2 |  -3.081889   .5880602    -5.24   0.000    -4.234904   -1.928874
    _Irace_3 |  -3.218312   .5551379    -5.80   0.000    -4.306776   -2.129848
    _Irace_4 |  -.2328243   1.109032    -0.21   0.834    -2.407315    1.941666
     _Isex_2 |  -3.643206   .3788777    -9.62   0.000    -4.386075   -2.900337
     _Isex_9 |  -3.820359   15.57913    -0.25   0.806    -34.36652     26.7258
       edyrs |   1.006814   .0766459    13.14   0.000     .8565331    1.157094
       _cons |   3.228584   1.158591     2.79   0.005     .9569229    5.500246
------------------------------------------------------------------------------

. outreg using mod1, se append

.
. xi: regress fhrwage i.race i.sex edyrs f_ed [w=wgt]
i.race            _Irace_1-4          (naturally coded; _Irace_1 omitted)
i.sex             _Isex_1-9           (naturally coded; _Isex_1 omitted)
(sum of wgt is   5.7244e+06)

      Source |       SS       df       MS              Number of obs =    3189
-------------+------------------------------           F(  7,  3181) =   74.37
       Model |  58264.8819     7  8323.55455           Prob > F      =  0.0000
    Residual |  356038.739  3181   111.92667           R-squared     =  0.1406
-------------+------------------------------           Adj R-squared =  0.1387
       Total |  414303.621  3188  129.957221           Root MSE      =   10.58

------------------------------------------------------------------------------
     fhrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Irace_2 |  -2.828454   .5854587    -4.83   0.000    -3.976369   -1.680539
    _Irace_3 |   -2.30416   .5687151    -4.05   0.000    -3.419246   -1.189075
    _Irace_4 |  -.4045806   1.102043    -0.37   0.714    -2.565368    1.756207
     _Isex_2 |  -3.610974   .3764164    -9.59   0.000    -4.349017    -2.87293
     _Isex_9 |  -4.906034   15.47749    -0.32   0.751    -35.25291    25.44084
       edyrs |   .6958816   .0896131     7.77   0.000     .5201762    .8715869
        f_ed |   1.168873   .1776414     6.58   0.000     .8205702    1.517177
       _cons |    2.27593   1.160037     1.96   0.050     .0014327    4.550427
------------------------------------------------------------------------------

. outreg using mod1, se append
```