**BEGIN LAB SESSION 1 HERE, WORKING YOUR WAY THROUGH THE CODE AND THEN WRITING CODE FOR YOUR COUNTRY’S TIME USE SURVEY TO ACCOMPLISH THE SAME TASKS:**

 **IDENTIFY PRODUCTIVE ACTIVITIES BY GROUP**

 **COMPUTE AGE-SEX AVERAGE TIME SPENT IN ACTIVITIES BY GROUP**

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
I. INTRODUCTION

ntta\_production.do
Gretchen Donehower, May 16, 2012

Description: Take dataset with American Time Use Survey (ATUS) data for 2009, with an observation for each activity reported by an ATUS respondent, and turn it into average time spent for each respondent in each activity being used to compute NTTA.

Note that this code is just an example based on the US survey!  Each country will need to make extensive changes based on the organization of the time use survey being used.

Code was written to be used with Stata Version 12.  To use the code with any other versions of stata, modifications may be necessary.
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
clear
set more off
use /data/lee0/gretchen/TIME\_USE/atusact\_2009.dta
keep tucaseid tuactdur24 tutier1code tutier2code tutier3co trtcctot\_ln trthh\_ln

/\*
SURVEY DETAILS:
Survey has one observation for each activity reported by an ATUS respondent. Activity durations are reported in minutes.  Each respondent was asked to account for all activities in one 24 hour period from 4am to 4am.  Activities are coded based on an activity coding scheme.  Weights are provided to create a
representative day of the year, correcting for days of the week, non-response, and sample design. Activity/duration data are in one file, weights and respondent characteristics are in another file.

Important variable descriptions:
  tucaseid: identifies ATUS respondent
  tuactdur24: variable giving time use, in minutes, of each activity; imputed for some
       respondents to make sure that total time accounted adds up to 24 hours (1440 minutes)
  tutier1code: part of ATUS activity coding scheme, identifies major activity groupings
  tutier2code: part of ATUS activity coding scheme, identifies activity sub-headings
  tutier3code: part of ATUS activity coding scheme, identifies lowest-level activities
  teage: age (single years of age from 15 to 79, 80 is ages 80-84, and 85 is 85+
  tesex: sex (1 for male, 2 for female)
  tufinlwgt: weight variable to be used with respondent-level data
  trtcctot\_ln: variable giving amount of time during which secondary childcare (age <13) was taking place
simultaneously (gives part of tuactdur24 that overlapped with secondary childcare)
  trthh\_ln: the part of trtcctot\_ln which was for own household children (difference trtcctot\_ln - trthh\_ln is secondary childcare for non-hh children)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
II. PRODUCTION
Generate "code" variable with a single numeric code for each activity:
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
tostring tutier1code, generate(tutier1)
replace tutier1="0"+tutier1 if length(tutier1)==1
tostring tutier2code, generate(tutier2)
replace tutier2="0"+tutier2 if length(tutier2)==1
tostring tutier3code, generate(tutier3)
replace tutier3="0"+tutier3 if length(tutier3)==1
gen CODE=tutier1+tutier2+tutier3
destring CODE,generate(code)
sort tucaseid

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Make variables for each of the 11 activities included in household production. Note that, for allocating consumption of time,  the "care" variables need to be split into separate variables based on the target of the care (inside the household or outside, child, elderly, or unspecified as in volunteering)
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
gen clean=tuactdur24\*(code==20101)
gen laund=tuactdur24\*(code==20102|code==20103)
gen cook=tuactdur24\*(code>=20201 & code<=20299)
gen hhmaint=tuactdur24\*((code>=20301 & code<=20499)|(code>=20701 & code<=20899))
gen lawngar=tuactdur24\*(code>=20501 & code<=20599)
gen hhmgmt=tuactdur24\*(code==20104|code==20199|(code>=20901 &
code<=29999)|(code>=160103 & code<=160108))
gen petcare=tuactdur24\*(code>=20601 & code<=20699)

/\*Category for purchasing goods and services has many code groups, so split up code on several lines. Includes all of "consumer purchases” (major category 07) and "purchasing household services" (major category 09) but only some of "purchasing professional & personal care services" (major category 08). Also has one category from government\*/
gen purch=0
replace purch=purch+tuactdur24\*(code>=70101 & code<=79999)
replace purch=purch+tuactdur24\*(code>=80101 & code<=80199)
replace purch=purch+tuactdur24\*(code>=80601 & code<=80799)
replace purch=purch+tuactdur24\*(code>=90101 & code<=99999)
replace purch=purch+tuactdur24\*(code==100103)

/\*Care variables divided by target of care.  "hh" ending for care within household, "nhh" for care outside of household but to targeted age group, "v" for volunteering where we don't know age target of
 care \*/
gen chcarehh=tuactdur24\*(code>=30101 & code<=30399)
gen chcarenhh=tuactdur24\*(code>=40101 & code<=40399)
gen elcarehh=tuactdur24\*(code>=30401 & code<=39999)
gen elcarenhh=tuactdur24\*(code>=40401 & code<=49999)
gen elcarev=tuactdur24\*(code>=150101 & code<=159999)
gen trav=tuactdur24\*((code>=180201 & code<=180499)|(code>=180701 & code<=180999))

/\*For comparative purposes, include variables for paid time and work related time (unpaid but related to paid work) \*/
gen paid=tuactdur24\*((code>=50101 & code<=50199)|(code>=50301 & code<=50399))
gen workrel=tuactdur24\*((code>=50201 & code<=50299)|(code>=50401 & code<=59999))

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Where secondary childcare is taking place if the primary activity is not work-related or household production, the entire time unit goes to childcare.  If the primary activity is  work-related or household production, give half of the unit to childcare and half of the unit to the primary activity.

Need to keep the secondary and primary childcare activities separate, though, because the target age groups are different (primary is age 0-17, secondary is age 0-12)

Just need to make sure that I don't double count a secondary childcare with a primary (double counting could occur if mix of household and nonhousehold children). In this case, privilege the primary activity.
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
gen sechh=trthh\_ln\*(trthh\_ln>0)
gen secnhh=trtcctot\_ln-sechh
foreach hhtype in hh nhh {
 foreach vvv in clean laund cook hhmaint lawngar hhmgmt petcare purch
elcarehh elcarenhh elcarev trav paid workrel {
   replace `vvv' = `vvv'-(sec`hhtype'/2) if `vvv'>0 & sec`hhtype'>0
   replace sec`hhtype'=sec`hhtype'/2 if `vvv'>0 & sec`hhtype'>0
 }
}
replace sechh=0 if chcarehh>0|chcarenhh>0
replace secnhh=0 if chcarehh>0|chcarenhh>0

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Collapse activity dataset to the respondent level (one observation for each respondent, variables represent total hours respondent spent in each activity on the interview day).
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
local hhproduction clean laund cook hhmaint lawngar hhmgmt petcare
purch chcarehh chcarenhh sechh secnhh elcarehh elcarenhh elcarev trav
collapse (sum) `hhproduction' paid workrel, by(tucaseid) fast
foreach vvv in `hhproduction' paid workrel {
 replace `vvv'=`vvv'/60
}
save profiles, replace

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Merge by respondent ID with files that have survey weights and respondent age and sex. Note that the file with the survey weights has one observation for each respondent just as the file just made above.  This merge is a one-to-one (1:1) merge and every respondent in "profiles.dta" has a match in "atusresp\_2009.dta".  For the next merge, the file "atusrost\_2009.dta" has one observation for each respondent and all respondent household members.  This is a one-to-many (1:m) merge and at this point we only want the data for the respondents, so we only keep those observations with a match (\_merge==3) in the second merge.
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
merge 1:1 tucaseid using /data/lee0/gretchen/TIME\_USE/atusresp\_2009.dta
drop \_merge
sort tucaseid tulineno
merge 1:m tucaseid tulineno using /data/lee0/gretchen/TIME\_USE/atusrost\_2009.dta
keep if \_merge==3
drop \_merge
keep teage tesex tufinlwgt tucaseid `hhproduction' paid workrel
/\* Next line is just to put the age for the 80-84 age group at the center so estimates look more accurate when plotted. \*/
replace teage=82 if teage==80

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Before collapsing to age profiles, save the microdata (individual-level) because that will be needed to calculate consumption and transfers.

Now collapse to (unsmoothed) age profiless, using survey weights to calculate accurate means.  Keep total survey weight data for each age group to use when smoothing.  Make one overall time summary variable.
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/
keep `hhproduction' paid workrel teage tesex tufinlwgt tucaseid
save microdata, replace

gen totwt=tufinlwgt
collapse (rawsum) totwt (mean) `hhproduction' paid workrel
[aw=tufinlwgt], by(teage tesex)
sort tesex teage
egen tothha=rowtotal(`hhproduction')
egen totwk=rowtotal(paid workrel)
sort tesex teage
save profiles, replace