Analyzing Poverty by Race with Stata

General Tips for Doing Microdata Stata Analysis

* **Use a Do file:** Save all your commands there so that you can replicate your analysis in the future. Include lots of comments so that you can understand why you did things a certain way.
* **Best way to learn is to practice:** Replicate a published American Fact Finder table. (You should come close, but no exact match since Public Use Microdata sample is 1/3 of the full sample.) Try out different Stata commands. Learn from the Stata code of other analysts.
* **Strategies to learn new Stata commands:**
  + Use the help Stata command. Type: help [name of command]
  + Use Stata menus. For example go to Data 🡪 Create or change data 🡪 Create new variable
  + Google Stata questions. There are many online tutorials and resources. UCLA has a very good website with instructional videos: <http://www.ats.ucla.edu/stat/stata/>
  + Reach out to us. One good way to start is by emailing all three of us a question: Vincent Palacios [palacios@cbpp.org](mailto:palacios@cbpp.org); Raheem Chaudhry [rchaudhry@cbpp.org](mailto:rchaudhry@cbpp.org); Danilo Trisi [trisi@cbpp.org](mailto:trisi@cbpp.org)
* **Use the Data Dictionary:** Reading the full data dictionary is a very good way to generate research ideas and become familiar with the type of analyses that you could do.

Description of Files for This Workshop Posted in Github

These are available at: <https://github.com/IMPACT2016/DataCorner>

|  |  |
| --- | --- |
|  | Description |
| ACS-Adv-DataDict14.pdf | ACS 2014 data dictionary |
| ACS-Adv-Handout\_V2.docx | Copy of this handout |
| ACS-Adv-PUMSDownloader.do | Stata Do file to download 1-year ACS microdata to your hard drive |
| ACS-Adv-StataDoFile-forWorkshop11-30-16.do | Stata Do file that I worked from today for this workshop (Contents of it are also pasted below.) |
| ACS-Adv-StataFile-2014pus\_FL-small.dta | ACS 2014 person data file for Florida in Stata format. That’s the data set I used for this workshop. |
| ACS-Adv-label\_ACS\_2014.do | Stata Do file that will add Census defined labels to the variables |
| ACS-Adv-Bonus-Stata-Code.do | Stata Do file that we used for a longer training. Includes more commands and other types of analyses that you can do. |

Example of an ACS family

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | RELP | AGEP | WAGP | SSIP | FINCP | POVPIP |
| Reference Person | 0 | 40 | 25,000 | 0 | 34,000 | 154 |
| Spouse | 1 | 42 | 0 | 5,000 | 34,000 | 154 |
| Child | 2 | 16 | 4,000 | 0 | 34,000 | 154 |
| Child | 2 | 13 | . |  | 34,000 | 154 |

Stata Code with Step-By-Step Instructions for Workshop Exercise

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\* 1. LOAD DATA

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

\*Load data file

use "C:\P\ACS\ACS2014pus\_FL.dta", clear

/\* Execute an existing DO file called "labelACS.do" that will add Census defined

\*\* labels to the variables \*/

do "C:\P\ACS\label\_ACS\_2014.do"

\*Check out the data

\*run some basic summary stats of key variables

sum SERIALNO ST RELP AGEP WAGP SSIP PINCP POVPIP RAC1P HISP

\*read variable labels

desc SERIALNO ST RELP AGEP WAGP SSIP PINCP POVPIP RAC1P HISP

label list relpLB

\*browse the data

browse SERIALNO ST RELP AGEP WAGP SSIP PINCP POVPIP RAC1P HISP

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\* 2. CREATE DEMOGRAPHIC VARIABLES

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

\*\* Create age category variable

generate byte AgeCat = 0

label variable AgeCat "3 Category Age Variable"

replace AgeCat = 1 if AGEP <= 17

replace AgeCat = 2 if AGEP >= 18 & AGEP <= 64

replace AgeCat = 3 if AGEP >= 65

label define AgeCatLB 1 "0-17" 2 "18-64" 3 "65+"

label values AgeCat AgeCatLB

\*Double check that it's working ok:

tab AgeCat

tab AGEP AgeCat

\*\* Create race category variable

generate byte RaceCat = 5 /\* Other: Not Hispanic, White, Black, or Asian \*/

label variable RaceCat "5 Category Race/Ethnic Variable"

replace RaceCat = 1 if RAC1P == 1 /\* White alone, Non-Hispanic \*/

replace RaceCat = 2 if RAC1P == 2 /\* Black alone , Non-Hispanic \*/

replace RaceCat = 4 if RAC1P == 6 /\* Asian alone , Non-Hispanic \*/

replace RaceCat = 3 if HISP >= 2 & HISP <= 24 /\* Hispanic \*/

label define RaceCatLB 1 "White Non-Hisp" 2 "Black Non-Hisp" 3 "Hispanic" 4 "Asian Non-Hisp" 5 "Other Non-Hisp"

label values RaceCat RaceCatLB

\*Double check that it's working ok:

tab RaceCat

tab RAC1P RaceCat

tab RAC1P RaceCat if HISP == 1

tab RAC1P RaceCat if HISP != 1

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\* 3. DEFINE POVERTY

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* Define who is in the poverty universe by excluding persons who

\*\* are unrelated children (persons under 15 who are not related to reference person)

\*\* -or- those that live in group quarters (dorms, nusring homes, youth homes, prisons,

\*\* mental facilities, etc.) and whose POVPIP value is missing. \*/

generate byte povuniv = 1

replace povuniv = 0 if AGEP < 15 & RELP >= 11

replace povuniv = 0 if RELP > 15 & POVPIP == .

label variable povuniv "Poverty Universe: Excludes some people living in group quarters and unrelated children under 15"

\*\* Create a binary variable that identifies poor families (below 100% of poverty).

generate byte inpoverty = (POVPIP < 100)

label variable inpoverty "Family income is below the poverty line"

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\* 4. PRODUCE TABLES

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

\*\* Show a table of population by state.

table ST [pweight=PWGTP], f(%15.0fc) row

\*\* Show a table of the population within the poverty universe by state.

table ST if povuniv==1 [pweight=PWGTP], f(%15.0fc) row

\*\* Show a table of population by age

table AgeCat [pweight=PWGTP] if ST == 12 & povuniv == 1, f(%15.0fc) row

\*\* Show poverty rates by age

table AgeCat [pweight=PWGTP] if ST == 12 & povuniv == 1, f(%15.3fc) c(mean inpoverty) row

\*\* Show population by age & race

table AgeCat RaceCat [pweight=PWGTP] if ST == 12 & povuniv == 1, f(%15.0fc) row col

\*\* Show poverty rates by age & race

table AgeCat RaceCat [pweight=PWGTP] if ST == 12 & povuniv == 1, f(%15.3fc) c(mean inpoverty) row col

\*\* Show number of people in poverty by age & race

table AgeCat RaceCat [pweight=PWGTP] if ST == 12 & povuniv == 1 & inpoverty == 1, f(%15.0fc) row col

Danilo Trisi

Center on Budget and Policy Priorities

trisi@cbpp.org

[www.cbpp.org](http://www.cbpp.org)