URSC 689 Coding Challenge 1 research log for Nathanael Rosenheim

February 11, 2020

How to make a function that prints the state name given the FIPS code?

When writing a program (and debugging it) have to remember to use program drop.

Error when program gets a FIPS code that does not have a state such as 3
Made an if statement to check this error

I wanted to make my program more "Full proof" by providing an error message. I know other programs do this. I tried to google search but did not find anything helpful
I looked at an existing do file
C:\ado\plus\a\asdoc.ado – using notepad++ to open the file… I found some helpful ideas for program writing.

Finished working on program
I can envision using the program to generate urls to download Census Data

February 10, 2020

Email instructions from Nathanael Rosenheim:
During class next week – we will have our first in-class coding demonstration.
The coding challenge is to make a program that lists the full names of all US States using a loop. Bonus points if you use a loop and an abstraction. To make the challenge competitive, I will give bonus points to the best code. Please have you code ready to share at the start of class.

Plan - Loops

Review Readings:
Long, J. S. (2009). The workflow of data analysis using Stata. College Station, TX: Stata Press.
https://www.stata.com/bookstore/workflow-data-analysis-stata/
Gentzkow, M., & Shapiro, J. M. (2014). Code and data for the social sciences: A practitioner's guide. University of Chicago mimeo.
https://people.stanford.edu/gentzkow/sites/default/files/codeanddata.pdf

Scott Long Loops – Look up in index – wow lots of entries on loops – this must be an important topic

Pages 92-106

Stata has foreach and forvalues

Stata Example Code

```
forvalues i = 40(5)80 {
```

```
}
```

Plan Abstraction

```
What is an abstraction?
- Not mentioned in Scott Long's book
  Gentzkow and Shapiro – "In programming, turning the specific instances
  of something into a general-purpose tool is known as abstraction." (p.
  24)
- Good for reducing redundancy (which reduces chances of error), code
  more readable

Abstraction example
program leaveout_mean
    syntax, invar(varname) outvar(name) byvar(varname)
    tempvar tot_invar count_invar
    egen `tot_invar'= total(`invar'), by(`byvar')
    egen `count_invar'= count(`invar'), by(`byvar')
    gen `outvar' = (`tot_invar' - `invar') / (`count_invar' - 1)
end

leaveout_mean, invar(pc_potato) outvar(leaveout_state_pc_potato)
byvar(state)
leaveout_mean, invar(pc_potato) outvar(leaveout_metro_pc_potato)
byvar(metro)
leaveout_mean, invar(hh_potato) outvar(leaveout_metro_hh_potato)
byvar(metro)

Scott Long also discusses programs on Page 111-119
```

Look at example code from previous programs

Example 1: SAS Program

P:\SVACS\Work\SVACS_0cv2_2010Census_2018-03-09.sas
Loops through list of states to run programs related to reading in Census Data. Census files use state abbreviations and names.

Example
https://www2.census.gov/acs2013_5yr/summaryfile/2009-2013_ACSSF_By_State_All_Tables/

www2.census.gov/acs2013_5yr/summaryfile/2009-2013_ACSSF_By_State_All_Tables/

United States Census Bureau

Name	Last modified	Size	Description
Parent Directory		-	
Alabama_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 08:25	117M	
Alabama_Tracts_Block_Groups_Only.zip	21-Nov-2014 08:32	55M	
Alaska_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 09:06	52M	
Alaska_Tracts_Block_Groups_Only.zip	21-Nov-2014 08:48	8.9M	
Arizona_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 08:57	73M	
Arizona_Tracts_Block_Groups_Only.zip	21-Nov-2014 08:33	72M	
Arkansas_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 09:13	125M	
Arkansas_Tracts_Block_Groups_Only.zip	21-Nov-2014 08:55	34M	
California_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 08:32	311M	
California_Tracts_Block_Groups_Only.zip	21-Nov-2014 09:31	407M	
Colorado_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 08:52	87M	
Colorado_Tracts_Block_Groups_Only.zip	21-Nov-2014 09:05	60M	
Connecticut_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 08:30	72M	
Connecticut_Tracts_Block_Groups_Only.zip	21-Nov-2014 09:05	41M	
Delaware_All_Geographies_Not_Tracts_Block_Groups.zip	21-Nov-2014 08:38	15M	

```
/*-------------------------------------------------------------------*/
/* Macro to call states                                              */
/*-------------------------------------------------------------------*/


%macro CallSt;
/*The CallSt macro is used to generate State 2 digit abbreviations see Appendix B
  of the technical documentation for a list of state codes                                    */
/*The Census summary file contains state 2 digit numeric codes from 1 to 72
  Note:  FIPS codes are NOT sequential so if a code does not exist such as 71
               The call execute statement will NOT run because there is no state abbrev    */
/*If you want just a single state, such as Alabama set the do statement to start
  and end at that state code such as %do i=1 %to 1; for Alabama                          */
%do i=&f_state %to &l_state;
      data _null_;
        stabbrv=compress(trim(lowcase(FIPSTATE(&i))));
        * Folder Is state name with _ for spaces;
        stname=compress(tranwrd(trim(FIPNAMEL(&i))," ","_"));
              /*Note:  DC and PR are not covered in FIPS state function, and the two digit
              function is not required to be 0 filled            */
          if &i=0  then stabbrv = 'us';
              if &i=11 then stabbrv = 'dc';
              if &i=72 then stabbrv = 'pr';
              /*FIPS Codes 60 and 66 are fpr American Samoa and Guam */
              if &i>56 and &i<72 then stabbrv = "--";
              /*If the function returns a state abbreviation then run the AllSeqs macro    */
              if stabbrv ^= "--" then do;
                      put stabbrv stname;
                      call execute('%Read2010Census(' || compress(stabbrv) ||','|| compress(stname) || ')');

              end;
      run;
%end;
%mend;
%CallSt;
```

Example 2: LODES Bulk Download

```
Program: LODES7_D0av1_bulkDownload_LEHDv72_2016-03-25.r

# What year do you want to download? e.g. _2013
```

```
# use _20 to download all years
year <- "_20"
# need a list of states with abbreviated state names (US postal abbrev.
# state name = STUSAB)
if (!file.exists("state.txt")){

download.file("http://web.archive.org/web/20141125122851/http://www.censu
s.gov/geo/reference/docs/state.txt", destfile = "state.txt")
} else {
  states <- read.table("state.txt", sep = "|", header = T, colClasses =
"character")
  states <- states[order(as.numeric(states$STATE)), ]
  states.keep <- states[-c(52:57),]
  nstates <- dim(states.keep)[1]
}
for (i in 1:51) {
  stusab <- tolower(states.keep[i, 2])
  fname <- paste0("lodes_", stusab)
```

Google Search:

"stata loop through states" – no immediate results examples of loops
Some help with
https://jearl.faculty.arizona.edu/sites/jearl.faculty.arizona.edu/files/Introduction%20to%20Loops%20in%20Stata.pdf
nothing stands out in google search

Plan – I need a list of all states

Google search – Census is a reputable source
https://www.census.gov/library/reference/code-lists/ansi.html

https://www2.census.gov/geo/docs/reference/state.txt?#
This file is available on the internet and has states and state equivalents

State and State Equivalents

National FIPS and GNIS Codes File

This file contains pipe delimited records for each state. The records are of the format:

FIPS State Code | Official United States Postal Service (USPS) Code | Name | Geographic Names Information System Identifier (GNISID)

For example:

23|ME|Maine|01779787

How to read in data – Template Workflow Example and Scott Long

```
insheet using "filename", clear

When I use insheet I only get 1 variable
help insheet
set delimiter to the pipe

Data  is in Stata – time to write a do file
Using do file template

G:\Team Drives\URSC689_S2020\TMPWF\Posted\00_templatedofile_2020-02-10.do

Saving work in WORKNPR Folder
G:\My
Drive\MyCourses\URSC689\WorkNPR\URSC689_01_IntroLoopsFunctions_2020-02-
10\
```

Data is read in – how do I loop over observations?

```
Google: stata loop through observations
Example code
https://www.stata.com/statalist/archive/2007-03/msg00525.html
```

```
sysuse auto, clear
gen domestic = .
local N = _N
forvalues i = 1/`N' {
        if foreign[`i'] == 0 {
                qui replace domestic = 1 in `i'
        }
        else {
                qui replace domestic = 0 in `i'
        }
}
```

How to make an abstraction (program) that returns state name given state fips?

```
Googled stata look up observation
Found
https://www.stata.com/statalist/archive/2008-08/msg00302.html
```

```
gen long obsn = _n
su obsn if name == "Jones", meanonly
local Jones_age = age[`r(min)']
```