```
. capture log close
```

LPO 9951: Tools of the Trade

Intro

This week we're going to get started with severa of the tools of the trade that we'll be using all year. The ideas I'm implementing here have been used by many analysts in the social sciences, but were captured best by Gentzkow and Shapiro in Code and Data for the Social Sciences: A Practitioner's Guide'. If you want to know why we're doing what we're doing this is an excellent resource.

Command Line

Having access to command line tools is very important when doing any kind of development. If you're on a Mac, the `terminal` program is used for command line interface with the computer. The best tool for a Windows computer is powershell. We'l l start out by making sure everyone has these.

Version Control

Version control is the general term for software that tracks changes in code (or other documents) and has resources for reverting or merging in changes. One of the most popular forms of version control is Git. We're going to start by using Github Deskto p, but later we'll switch to interfacing with git and github via the command line.

You'll need to sign up for a github account at (https://github.com/).

Then you'll want to download github desktop.

The first thing I want you to do is to create a clone of our class directory. Go to (https://github.com/wdoyle42/lpo_prac), copy the url, then go to File–> Clone directory, then click the "URL" tab, and paste in the URL.

At a minimum, before every class you'll want to sync your directory with my changes.

Next, I want you to create a private repository that contains your work for this class. To do this, start in github by creating a repository, and in the repository, create a brief readme that states that this is the working directory for your practicum course. Next, add me (wdoyle42) as a collaborator on this repostiory. Clone this repostiory to your computer, create a directory for assignments, and then create a do file in the assignments directory named 02-assignment.do. Commit and push this empty d o file to the repository. I'll double check that I have access to it.

Markdoc

Markdoc is a literate programming package for Stata. Literate programming is a (pretty old) idea that has been VERY slow to catch on among social scientists, but means combining our writing and our code into a sing le document.

To install markdoc, go to Stata and first install the `github` package:

net install github, from("https://haghish.github.io/github/")

Next, install the markdoc package:

```
github install haghish/markdoc, stable
```

Pandoc

To take full advantage of markdoc, we need two additional tools: pandoc, and a Latex installation. Pandoc is a univeral document translator. Download it here: (https://github.com/jgm/pandoc/releases/tag/2.14.2). Once you've downloaded it and installed i t, you can check on the installation in your terminal using

```
pandoc --version
```

Latex

Latex is a typesetting program that has a huge number of useful features for technical writing. We won't author documents in latex for this class, but we will use its functionality. To download Latex, go here: (https://www.latex-project.org/get/)

Helpful template and workfolow ideas

From Matt Ingram

Running your first markdoc documents

Create another directory for lessons in your github repository. Copy today's lesson into that directory. In Stata, run

```
markdoc tools_of_the_trade.do, export(md)
```

This will create a markdown document. Markdown is a simple syntax for generating html, and it serves as a great "source" language for a variety of typesetting programs, including Latex and Word.

```
. clear all                                  // clear memory


. set more off                               // turn off annoying "__more__" feature

. log using "tools_of_the_trade.log", replace
```

```
-------------------------------------------------------------------------------
      name:  <unnamed>
       log:  /Users/doylewr/lpo_prac/lessons/s1-02-stata_basics/tools_of_the_trade.log
  log type:  text
 opened on:   1 Sep 2021, 10:56:14

. net search renvars
(contacting http://www.stata.com)

4 packages found (Stata Journal and STB listed first)
---------------------------------------------------

dm88_1 from http://www.stata-journal.com/software/sj5-4
    SJ5-4 dm88_1.  Update:  Renaming variables, multiply and... / Update:
    Renaming variables, multiply and systematically / by Nicholas J. Cox,
    Durham University, UK / Jeroen Weesie, Utrecht University, Netherlands /
    Support:  n.j.cox@durham.ac.uk, j.weesie@fss.uu.nl / After installation,

dm88 from http://www.stata.com/stb/stb60
    STB-60 dm88.  Renaming variables, multiply and systematically / STB insert
    by Nicholas J. Cox, University of Durham, UK / Jeroen Weesie, Utrecht
    University, Netherlands / Support:  n.j.cox@durham.ac.uk
    j.weesie@fss.uu.nl / After installation, see help renvars

cleanchars from http://fmwww.bc.edu/RePEc/bocode/c
    {c 39}CLEANCHARS{c 39}: module to replace specific characters or strings in variab
    names and/or variable labels and/or string variable values and/or value
    label names and levels with stated characters/strings (using 1-1 or m-1
    match) / cleanchars is a program that helps out with replacing /

renvarlab from http://fmwww.bc.edu/RePEc/bocode/r
    {c 39}RENVARLAB{c 39}: module to rename variables, with option of using variable
    labels to create new variable names / This command is an extension of
    renvars (also available from / SSC), which renames a list of variables by
    applying the given / transformation to all of the variables. It has all of


. webuse school, clear

. save school, replace
file school.dta saved

. outsheet using "school_data.csv", comma replace

. insheet using "school_data.csv", comma clear
(11 vars, 95 obs)
```

```
. outsheet using "school_data.tsv", replace

. insheet using "school_data.tsv", clear
(11 vars, 95 obs)

. describe

Contains data
 Observations:            95
    Variables:            11
-------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------
obs             byte    %8.0g
pub12           byte    %8.0g
pub34           byte    %8.0g
pub5            byte    %8.0g
private         byte    %8.0g
years           byte    %8.0g
school          byte    %8.0g
loginc          float   %9.0g
logptax         float   %9.0g
vote            byte    %8.0g
logeduc         float   %9.0g
-------------------------------------------------------------------------------
Sorted by:
     Note: Dataset has changed since last saved.

. label data "Voting on school expenditures"

. label variable loginc "Log of income"

. label variable vote "Voted for public school funding"

. describe

Contains data
 Observations:            95                  Voting on school expenditures
    Variables:            11
-------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------
obs             byte    %8.0g
```

```
pub12           byte    %8.0g
pub34           byte    %8.0g
pub5            byte    %8.0g
private         byte    %8.0g
years           byte    %8.0g
school          byte    %8.0g
loginc          float   %9.0g                 Log of income
logptax         float   %9.0g
vote            byte    %8.0g                 Voted for public school funding
logeduc         float   %9.0g
-------------------------------------------------------------------------------
Sorted by:
     Note: Dataset has changed since last saved.

. tab vote

  Voted for |
     public |
     school |
    funding |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         36       37.89       37.89
          1 |         59       62.11      100.00
------------+-----------------------------------
      Total |         95      100.00

. label define voteopts 0 "no" 1 "yes"

. label values vote voteopts

. tab vote

  Voted for |
     public |
     school |
    funding |      Freq.     Percent        Cum.
------------+-----------------------------------
         no |         36       37.89       37.89
        yes |         59       62.11      100.00
------------+-----------------------------------
      Total |         95      100.00

. gen inc = exp(loginc)

. sum loginc inc
```

```
    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
      loginc |         95    9.971017    .4118853      8.294      10.82
         inc |         95    23093.31     8871.35     3999.8   50011.07

. sum inc

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
         inc |         95    23093.31     8871.35     3999.8   50011.07

. gen inc_bin = 0

. replace inc_bin = 1 if inc > r(mean)
(30 real changes made)

. egen inc_q = cut(inc), group(4)

. recode inc_q (0 = 1 "First Quartile") ///
     (1 = 2 "2nd Quartile") ///
     (2 = 3 "3rd Quartile") ///
     (3 = 4 "4th Quartile"), gen(new_inc_q)

(95 differences between inc_q and new_inc_q)

. gen ptax = exp(logptax)

. gen taxrate = ptax / inc

. log close                                 // close log
      name:  <unnamed>
       log:  /Users/doylewr/lpo_prac/lessons/s1-02-stata_basics/tools_of_the_trade.log
  log type:  text
 closed on:   1 Sep 2021, 10:56:15
-------------------------------------------------------------------------------

. exit                                      // exit script
```