SOCIAL SCIENCES RESEARCH CENTER

THE UNIVERSITY OF CHICAGO **DIVISION OF THE SOCIAL SCIENCES**

Social Science Computational Resources Series

February 18, 2020

Brooke Luetgert, PhD

About this Series

This five part series is designed to expose participants to working with data in various software environments. While many tools will "get the job done", knowing the tips and tricks of multiple alternatives helps the user to work more efficiently and confidently through data analysis and presentation. Our focus is on good data practice, documentation and reproducibility across software platforms.

This is an introductory series for absolute beginners!

Contact Details

Brooke Luetgert, Computational Scientist at RCC

Email: luetgert@uchicago.edu
Office: TAAC 2, 5607 South Drexel
Telephone: (773)-834-5313

RCC Help Desk: Reg. 216, Mon-Fri 9AM-5PM

Materials on GitHub- use search bar, enter user:luetgert
Our folder is **luetgert/SSD_RCC_WorkshopSeries**

Plan for Today

Our focus will be two-fold:

1)We want to log on to Acropolis and access Stata

2)We want to work our way through a few basic regression and plotting tasks, then work on uploading and downloading data from Acropolis to our local machine.

R or Python will not be addressed today.

Connect with Acropolis

Stata is available in the SSD on Acropolis. You can download the EasyVNC client from https://sscs.uchicago.edu/2019/07/18/how-to-run-easyvnc/ .

Users can apply for accounts on Acropolis at https://iota.src.uchicago.edu/ .

Information about file transfers can be found here https://sscs.uchicago.edu/2019/07/17/how-do-i-access-files-on-the-server/.

Download Findings- Windows

open a File Explorer window (such as by clicking on the Computer icon), and type in a double backslash **** immediately followed by the address of the server and then Enter key. For **Acropolis** : **\\sscs-fs0.uchicago.edu**

Supply your CNET ID username and password. You must use the prefix **ADLOCAL**
For example, if your CNET ID is johnsmith, you would use **ADLOCAL\johnsmith** as your username. You should see your files.

Download Findings- Mac

Open a Finder Window, then click on the **Go** menu item, then click on **Connect to Server . . .** In the address bar that appears, type
smb://sscs-fs0.uchicago.edu
and then click the **Connect** button.

Supply your CNET ID username and password.
You must use the prefix **ADLOCAL**
For example, if your CNET ID is johnsmith, you would use **ADLOCAL\johnsmith** as your username. You should see your files.

Which Stata?

Stata is available in several versions: Stata/IC (the standard version), Stata/SE (an extended version) and Stata/MP (for multiprocessing). The major difference between the versions is the number of variables allowed in memory, which is limited to 2,047 in standard Stata/IC, but can be much larger in Stata/SE or Stata/MP. The number of observations in any version is limited only by memory.

Why Stata?

Stata requires that the entire dataset to be analyzed must reside in memory. This brings a considerable speed advantage, but implies that you may need more RAM (memory) on your computer.

Stata has traditionally been a command-line-driven package that operates in a graphical (windowed) environment. Stata version 11 (released June 2009) contained the first graphical user interface (GUI) for command entry via menus and dialogs.

Reproducibility

Stata makes reproducibility very easy through a log facility, the ability to generate a command log (containing only the commands you have entered), and the do-file editor which allows you to easily enter, execute and save sequences of commands, or program fragments.

Save the do-file to use it as the starting point for a similar set of data management or statistical operations.

>> Enter data early, record changes and models

Newest Features

Stata has made huge changes to the do-file editor. The editor can now deal with legacy code, auto-indent code to increase legibility, visually mark wrapped lines. Functions similar to Markdown.

New command **putdocx** lets you create Word (.docx) files more easily. You can create documents with

- formatted paragraphs
- tables
- embedded Stata graphs (or any PNG, JPEG, EMF, or TIFF files)