

Introduction to FarmShare for Statistical Packages (Windows)

This guide is intended to help first-time users understand how to remotely access statistical software programs on Stanford’s FarmShare from a PC.

FarmShare is one of Stanford’s shared research computing environments. It allows users to access statistical and other software programs at no cost on their own computers through the internet. It also connects them to Stanford’s powerful computers for much faster processing of computing-intensive tasks than would be possible on single desktop or laptop.

To use FarmShare, you must have a SUNet ID (Stanford University Network Identifier). If you do not have a SUNetID, visit www.stanford.edu/services/sunetid for more information on obtaining one.

Table of Contents

<i>Introduction to FarmShare for Statistical Packages (Windows)</i>	1
<i>Quick Start: Accessing Software in FarmShare</i>	2
<i>Introduction</i>	2
<i>Installation of Software for Using FarmShare</i>	3
<i>FarmShare Environments</i>	4
<i>Accessing Files</i>	4
<i>Two Methods of Using FarmShare</i>	5
<i>(1) Accessing FarmShare in Interactive Mode</i>	5
<i>(2) Accessing FarmShare in Batch Mode</i>	9
<i>For More Information and Assistance</i>	11
<i>SSDS Software Services at Stanford</i>	12

Quick Start: Accessing Software in FarmShare

As an overview, the steps to using software (in this example, Stata) are:

1. Download and install SecureCRT/SecureFX, and mobaXterm
2. Upload files you want to access in the software to your AFS space (afs.stanford.edu)
3. Open mobaXterm. Click on “**Sessions**” and select “**New session**”. Choose the SSH option.
4. Fill out the “remote host” field with **rice.stanford.edu** and type your *SUNet ID* in the “specify username” field. Type your password and 2-step authenticate.
5. Type the following commands in mobaXterm to begin Stata:
 - module load stata-se/15
 - xstata

The document below provides more detailed explanations and steps.

Introduction

Why would I use FarmShare for statistical packages?

- You want to access statistical software for *free*.
- You want to access statistical software *remotely*, rather than from your dorm or library computer cluster.
- Your dataset includes too many variables to open in your Stata IC (you need Stata SE/MP) on your personal computer. You need a *more powerful version of statistical software*.
- The tasks you need to perform are *computationally intensive* and take up all of your person computer’s resources (leaving you unable to watch that silly cat video your mom sent you while also running complex algorithms in Stata).

In this document, words appearing in **bold type** are commands in Unix or a statistical software package and should be typed as shown. Conversely, words that are *italicized* are not supposed to be typed verbatim; instead, you are to substitute another word for them. For example, when you see *filename* in this document, you should substitute the actual filename you would like your file to have. It is also important to remember that Stata, R, and Unix are case sensitive, though SAS is not. In Stata, R, and Unix, name1, Name1, and NAME1 are all different.

Through this document all examples will assume that you are on a rice machine in the Farmshare system since rice is recommended for most users and, as explained later, users access wheat and corn through rice. Each line on your terminal program will start with the name of the machine you are remotely logged into, followed by a prompt character such as “>”. In the document, we

use “rice02>” for the computer and prompt when giving examples of commands to type into FarmShare directly. When giving examples to type into a program (e.g., SAS, R) accessed through FarmShare, we use “>” alone as the prompt. Do not type these characters when entering your commands at the command line.

Installation of Software for Using FarmShare

Before making use of FarmShare, you will need to download a number of software programs that are available for free to members of the Stanford community.

In addition to giving information on each available software program, the above webpage also contains links for downloading the desired programs. The specific programs you will need to install depend on how and where you will be using FarmShare. For example, Windows users who wish to access a graphical user interface (GUI) version of Stata will need to download and install an X terminal such as mobaXterm whereas those who only wish to submit code and view the output in a file will need the program SecureCRT. The following list gives the programs that are needed for FarmShare or that make using FarmShare more convenient. You also may need Kerberos for Windows or Cisco AnyConnect to authenticate off-campus, see the **Authenticating Off-Campus** guide, available here: <https://ssds.stanford.edu/software-resources/getting-started-guides-documents>

Software for Working in FarmShare			
Software Package	Importance	Purpose	For More Information/Downloads
SecureCRT/ SecureFX (installer file installs both programs)	Needed for submitting code	SecureCRT allows users to login to FarmShare and encrypts the session. SecureFX is optional and allows users to view the files and folders for use in FarmShare.	Essential Stanford Software (itservices.stanford.edu/service/ess)
mobaXterm	Needed for using Graphical User Interface	Provides access to programs with Graphical User Interface in FarmShare	mobaXterm (http://mobaxterm.mobatek.net/)
Stanford Open AFS for Windows	Optional (access to afs.stanford.edu is sufficient)	Makes accessing files for use in FarmShare easier	Essential Stanford Software (itservices.stanford.edu/service/ess)

FarmShare Environments

When accessing FarmShare, users have a choice of which environment to use with different levels of computing power and different capabilities. Cardinal, which has the least amount of computing power of the FarmShare environments, is for low-intensity tasks such as checking email and so is not recommended for analyzing data.

Rice is likely the most commonly used for data analysis. It has the advantage of offering a good amount of computing power but also allows users to access a larger range of statistical software packages (SAS, Stata, R) than is possible in Wheat or Oat. It also works for interactive (where users type and submit one line of syntax at a time and view the results as they go) as well as non-interactive jobs (where users type and submit all syntax at once). Additionally, rice offers an x-simulator through X-Windows, which features a graphical user interface among others for Stata, SAS, and R users. The graphical user interface versions of statistical software programs have multiple windows and so have a similar feel to how the programs operate on a personal computer or a Mac.

For more information on FarmShare environments:

<https://itservices.stanford.edu/service/sharedcomputing/environments>

Accessing Files

Because FarmShare is not on your local computer, you need a method to access your files while in the FarmShare environment and then retrieve finished files from the FarmShare session to bring to your local computer. The simplest approach is to use Stanford's AFS (Andrew File System) space. AFS is a file system that enables efficient file sharing between clients and servers. AFS files are accessible via the Web (afs.stanford.edu) or through file transfer programs such as OpenAFS (see SSDS guide *Introduction to AFS* for more details).

The first step in using FarmShare in any capacity is to copy your relevant files (datasets, command scripts, etc.) into your AFS space by going to afs.stanford.edu and using the appropriate menu items to copy or upload your files.

Two Methods of Using FarmShare

There are two primary ways of interacting with FarmShare:

- (1) Interactively through an X-simulator or
- (2) Submitting code as “batch jobs”.

Using software interactively through X enables the use of statistical program’s familiar Graphical User Interface (GUI) to run new analyses and view output in real time. For non-computationally-intensive tasks, this is the most user-friendly.

Submitting code as “batch jobs” for statistical software involves users submitting the full script (e.g., Stata’s .do files) of the entire analysis in advance, and after submitting the “job” to FarmShare’s servers, users wait for the script to conclude running and view output files. This method of using FarmShare is useful for computationally-intensive tasks for the ability to take advantage of parallel processing and not being required to stay connected to FarmShare after submitting the job code.

(1) Accessing FarmShare in Interactive Mode

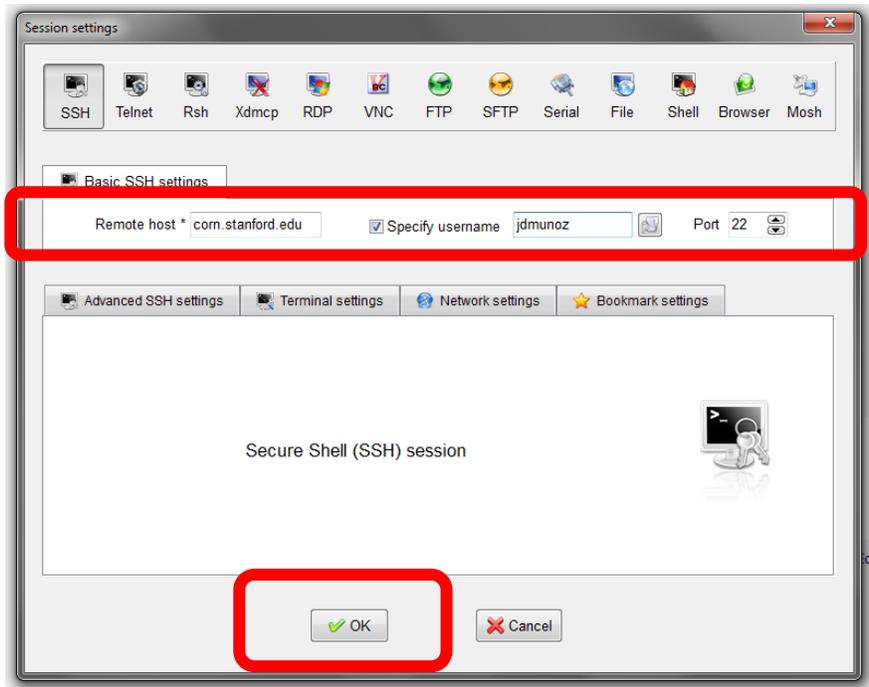
Logging in to FarmShare

There are two separate ways of accessing FarmShare. The first way we will cover here is accessing FarmShare through X-Windows. As mentioned above, access to FarmShare via X-Windows is needed if you want to use a version of SAS, Stata, or R with a graphical user interface. If you do not need to use the graphical user interface, SecureCRT is sufficient. We detail the procedure using the x-simulator [mobaXterm](#) below.

Download and install [mobaXterm](#). Once installed, use the “Sessions” tab to create a “New session”.

Then, choose the SSH session option.

When prompted, fill out the “Remote host” field with “rice.stanford.edu” and “Specify username” with your SUNet ID and click **OK**.



You will then be prompted for your SUNet password. (When typing your password, text will not appear in the terminal window. The connection is encrypted and secure.)

After two-step authenticating, you will enter Rice with graphical user interface capabilities. See the section *Loading Software into a FarmShare Session* to continue loading and beginning your software.

Loading Software into a FarmShare Session

Once in FarmShare, you will want to access the statistical software program of your choice. For Stata and SAS, which are not open source programs like R is, you will need to start by loading the program. The first step to loading a program is to type

```
rice02> module avail
```

at the FarmShare prompt and hit enter on the keyboard. You should see a list of available programs appear on your screen. Locate the program you would like to use and then type

```
rice02> module load program_name
```

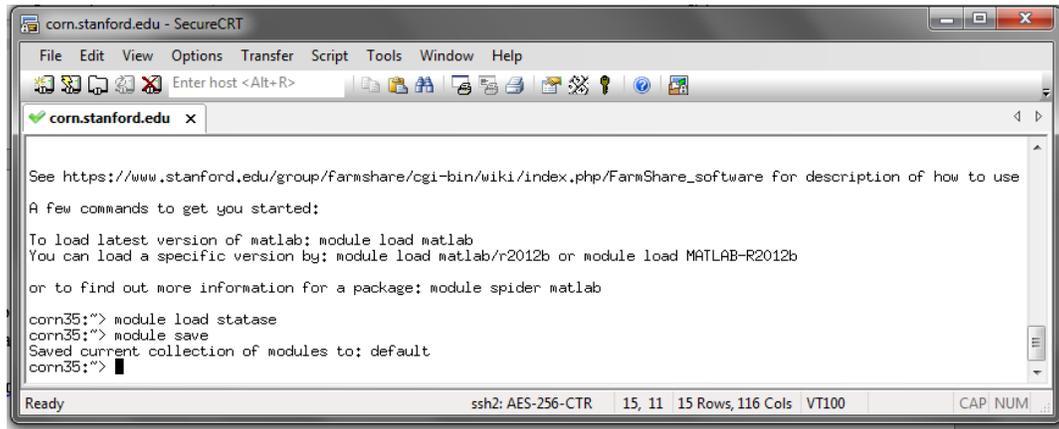
where *program_name* is the name of the desired program exactly as it appears. Then hit enter. Note that the FarmShare environment is case sensitive, so you will need to use the same upper-case and lower-case letters in the name of the program as it appears on your screen.

To pull up a list of programs that you have loaded into your FarmShare session, type

rice02> **module list**

and hit enter. To have a set of modules loaded automatically every time you start a session, load the modules you want to use and type

rice02> **module save**

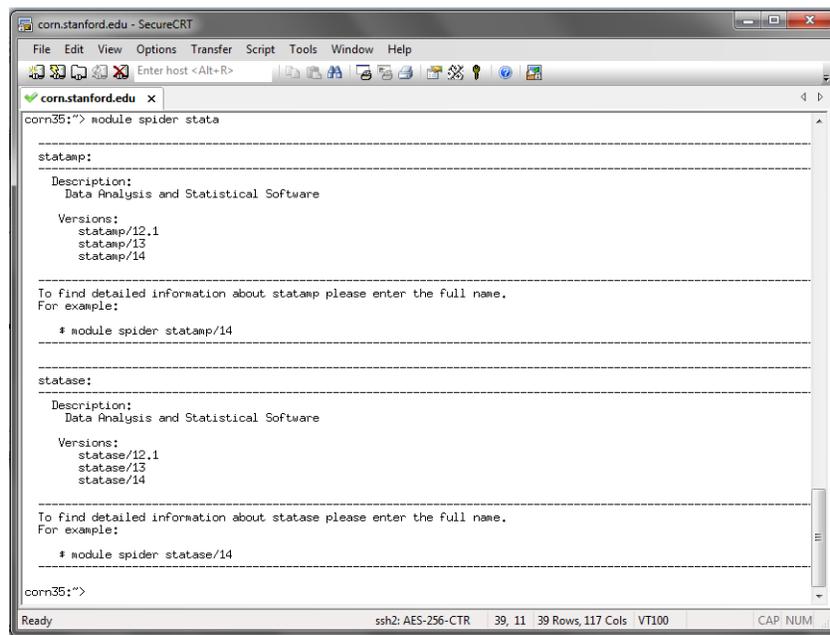


```
corn.stanford.edu - SecureCRT
File Edit View Options Transfer Script Tools Window Help
Enter host <Alt+R>
corn.stanford.edu x
See https://www.stanford.edu/group/Farmshare/cgi-bin/wiki/index.php/FarmShare_software for description of how to use
A few commands to get you started:
To load latest version of matlab: module load matlab
You can load a specific version by: module load matlab/r2012b or module load MATLAB-R2012b
or to find out more information for a package: module spider matlab
corn35:> module load stata
corn35:> module save
Saved current collection of modules to: default
corn35:>
```

The modules you save will be automatically loaded into your future sessions. If some modules are still unavailable in your next session, type **module update**.

To look for a particular version of a program use ‘**module spider program**’ and load the software version you want to use. For example, to see all the available versions of Stata type:

rice02> **module spider stata**



```
corn.stanford.edu - SecureCRT
File Edit View Options Transfer Script Tools Window Help
Enter host <Alt+R>
corn.stanford.edu x
corn35:> module spider stata
-----
statamp:
Description:
Data Analysis and Statistical Software
Versions:
statamp/12.1
statamp/13
statamp/14
-----
To find detailed information about statamp please enter the full name.
For example:
# module spider statamp/14
-----
stataase:
Description:
Data Analysis and Statistical Software
Versions:
stataase/12.1
stataase/13
stataase/14
-----
To find detailed information about stataase please enter the full name.
For example:
# module spider stataase/14
-----
corn35:>
```

To unload all the modules loaded into your session type:

```
rice02> module purge
```

To use a program in interactive mode once you have loaded it following the instructions above, type

```
rice02> program_name
```

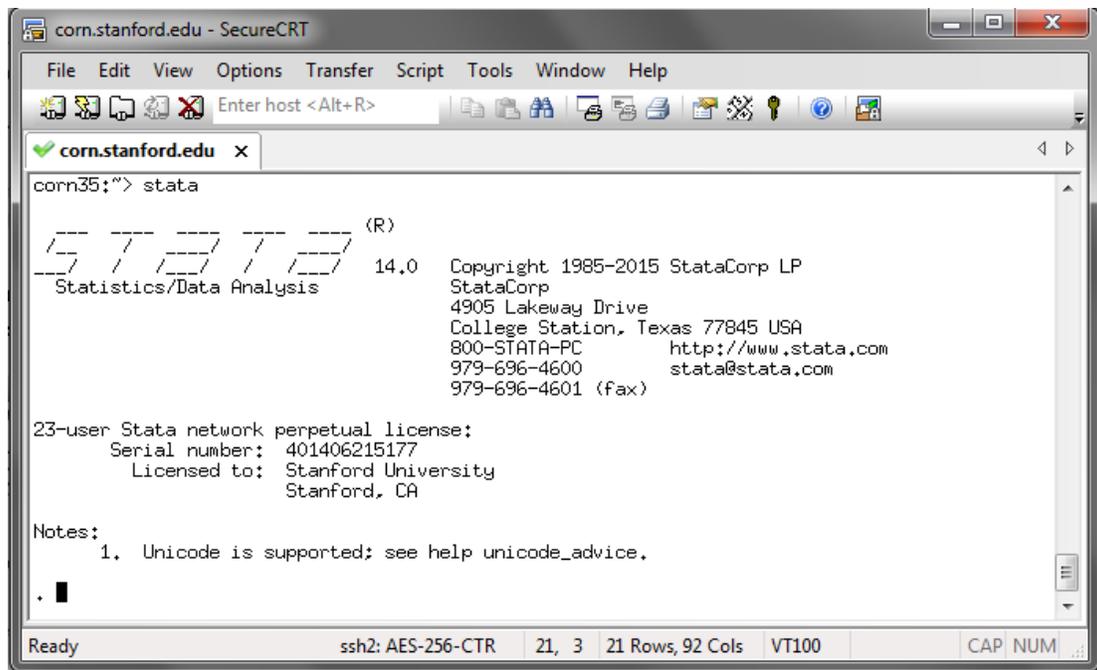
For example, the process for loading the most up to date Stata SE in FarmShare would include the following commands

```
rice02> module load stata-se/15
```

Once your software is loaded, there are two ways to start using the program.

(A) To use without the graphical interface, type:

```
rice02> stata
```



```
corn.stanford.edu - SecureCRT
File Edit View Options Transfer Script Tools Window Help
Enter host <Alt+R>
corn.stanford.edu x
corn35:~> stata
(R)
-----
Statistics/Data Analysis 14.0 Copyright 1985-2015 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)

23-user Stata network perpetual license:
Serial number: 401406215177
Licensed to: Stanford University
Stanford, CA

Notes:
1. Unicode is supported; see help unicode_advice.

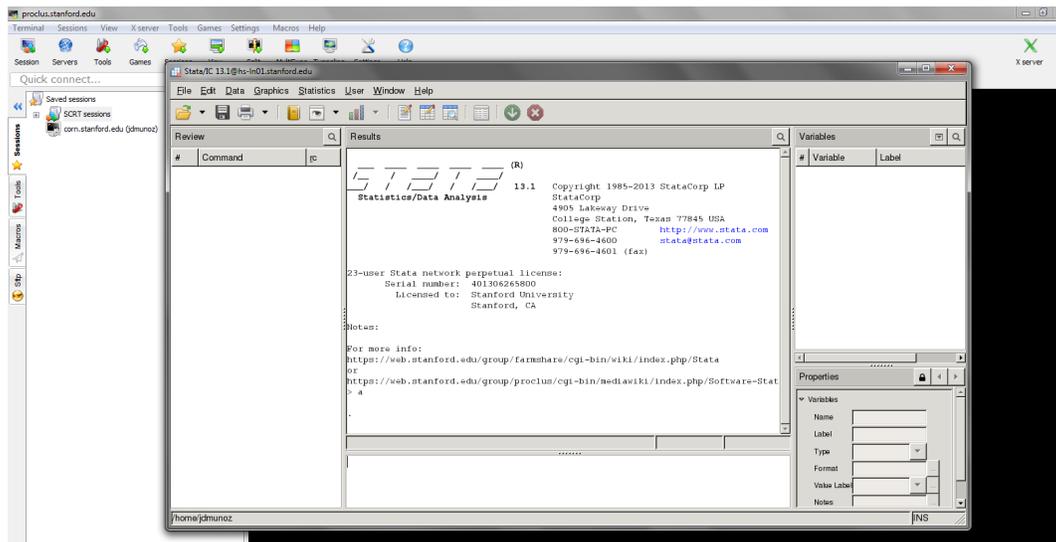
Ready ssh2: AES-256-CTR 21, 3 21 Rows, 92 Cols VT100 CAP NUM
```

Once a program is started, you can type software-specific commands directly on the command prompt. Note that, for most of the programs, the default working directory will be your AFS space at Stanford (see *Accessing Files* section above). That is, the software you are using will be searching for files and exporting output from and to your AFS space. Make sure that all the files and scripts for your session are uploaded into your AFS folder, which you can access at afs.stanford.edu. You can also change the working directory using the corresponding commands for each software.

(B) Or, to use the interactive version with graphical user interface, type:

rice02> *xstata*

When using Stata with the GUI, your end result after loading the module and beginning Stata, will look like this:



The drop-down menus and all other graphical user interface contents are shown on your screen through the x-simulator.

Type **exit** to close your program or type **quit()** to exit R. To end your session on FarmShare, also type **exit**.

(2) Accessing FarmShare in Batch Mode

In addition to accessing FarmShare through X-Windows, you may have more computationally-intensive tasks that you would like to run from start to finish without doing so in real time.

To use statistical software programs in batch mode, you will need to type all of the commands you wish to run in the program into a text editor (or the program's command line coding file) and then save the commands as a file that you will then submit to the program.

Submitting files with code to statistical software programs

The following table gives the Slurm syntax to submit a file named *filename* to R, SAS, and Stata. Note that *filename* must not contain spaces (so name the file *file_name* rather than *file name*). Type the specified syntax at the FarmShare prompt when you are in the folder/directory in FarmShare that contains the file you wish to submit. (See the section below titled **Common**

Slurm Commands for directions to change folders/directories in FarmShare.) For the syntax below to work for SAS and Stata, you will need to have loaded the needed program into your FarmShare session. (See the section above titled *Loading Software into a FarmShare Session*.) You will also need to load R into your FarmShare session if you wish to use the newest version available (v3.3.0) but not if you would like to use the default version of R (v3.0.2).

Program	Syntax	Notes
SAS	rice02> sas -memsize 0 filename &	The memsize option allows SAS to use the amount of memory needed for the job, even if this amount is greater than the default memory size allowed for SAS jobs in the system.
Stata	rice02> stata -b do filename &	This syntax assumes that <i>filename</i> has the extension .do ; when creating new files with a text editor, make sure to give them the .do extension in the filename (e.g., <i>filename.do</i>).
R	rice02> R --vanilla --no-save < filename	The --vanilla part of the syntax specifies that R will be used in batch mode instead of interactive mode.

To loop through a program’s command line coding file multiple times, you will also need to create a “.sbatch” file, which tell servers to open “module” (program) and run script (e.g., .do file) on dataset multiple times. You can tell servers to execute submit file with code:

```
rice02> sbatch example.sbatch
```

To loop through submitting this X times, enter:

```
rice02> for i in {1..X}; do  
  
          sbatch example.submit  
  
          done
```

Common Slurm Commands

This section focuses on basic commands in Slurm to help you navigate once in FarmShare, and the subsequent section will go over how to access statistical software packages in batch and interactive mode in FarmShare.

Once you are in FarmShare, you will need to use some Slurm commands to work with files and folders where you store your work. Files, such as data files, that you need to access in your FarmShare session should be saved in your AFS space. When you first connect to FarmShare, you will be in your home directory, or folder, in your AFS space and will have access to the files stored there. The commands below allow you to switch folders and manage your files. To use the commands, type them as they appear, substituting relevant folder and file names, and then hit enter on your keyboard. Note that Slurm is case sensitive, so you must type the name of folders and files exactly as they are. Both the commands and their results are listed in the table below.

Slurm commands for working with folders and files in FarmShare	
Command	Result
rice02> cd " <i>folder name</i> "	Navigates to " <i>folder name</i> " and grants access to the files it contains
rice02> cd ~	Navigates to home directory (the home folder in your AFS space)
rice02> ls	Pulls up a list of files in the current folder
rice02> ls -l	Pulls up a list of files in the current folder with details such as the files' dates
rice02> cp " <i>original filename</i> " " <i>copied filename</i> "	Creates a copy of " <i>original filename</i> " and gives it the name " <i>copied filename</i> "
rice02> rm " <i>filename</i> "	Deletes " <i>filename</i> "

Note: The commands above to navigate to folders are for the rice session within FarmShare. If applicable, you will also need to specify within the statistical software package that you use (SAS, R, or Stata) which directory or folder you would like to work in.

If you are new to working with your AFS, you can read more by clicking on the link for Introduction to AFS on this page: <https://ssds.stanford.edu/software-resources/getting-started-guides-documents>

For More Information and Assistance

The IT services at Stanford have a collection of webpages that provide information on FarmShare, including what it is and how to access it. You can find them here:

<https://itservices.stanford.edu/service/sharedcomputing>

The FarmShare wiki page is available at (including tutorial):

<https://srcc.stanford.edu/farmshare2>

You can get e-mail support at research-computing-support@stanford.edu

You can also get help from the Stat, Math, Algorithmic and Computational Consulting (SMACC) group at Stanford here: <http://stanford.edu/~rezab/smacc/>

Accessing FarmShare on a Mac

To access software on a Mac, see the document *Introduction to FarmShare for Statistical Packages (Mac)*.

SSDS Software Services at Stanford

If you have questions about using statistical software packages in a FarmShare environment, please contact the software consultants at Social Science Data and Software. Our website is <http://ssds.stanford.edu>. The software consultants are available during the academic year on a walk-in basis. Please see our website for current walk-in hours.

Copyright © 2018, by The Board of Trustees of the Leland Stanford Junior University. Permission granted to copy for non-commercial purposes, provided we receive acknowledgment and a copy of the document in which our material appears. No right is granted to quote from or use any material in this document for purposes of promoting any product or service.

Software Support, Social Science Data and Software

Document revised: 9/8/18