

# OES Visiting Researcher Guide

Revised 10/11/18

## Contents

<b>Administrative Information .....</b>	<b>2</b>
<b>Points of Contact .....</b>	<b>2</b>
<b>Data Access .....</b>	<b>2</b>
<b>Onsite Appointments .....</b>	<b>2</b>
<b>Passwords .....</b>	<b>2</b>
<b>Research Server .....</b>	<b>3</b>
<b>Research Server Uploads .....</b>	<b>3</b>
<b>How to Transfer Files using Tectia Secure File Transfer .....</b>	<b>3</b>
<b>Graphical User Interface (GUI) versions of SAS and STATA .....</b>	<b>4</b>
<b>Navigating Graphical User Interface (GUI) of SAS .....</b>	<b>5</b>
<b>Information for Outside Researchers Using OES Microdata .....</b>	<b>7</b>
<b>Non-Disclosure Review .....</b>	<b>8</b>
<b>Project Amendments .....</b>	<b>9</b>
<b>Project Extensions .....</b>	<b>9</b>
<b>Transferring Sponsoring Institutions .....</b>	<b>9</b>
<b>Safety .....</b>	<b>10</b>
<b>Building Evacuation .....</b>	<b>10</b>
<b>Shelter In Place .....</b>	<b>10</b>

## Administrative Information

### Points of Contact

#### Project Coordinators:

- Michael Soloy (soloy.michael@bls.gov; 202.691.6510)
- The OES Project Coordinators are responsible for maintaining researcher' access to OES data, and arranging for both the uploading of files and non-disclosure review of output.

#### Office of Administration:

- Erin Good (good.erin@bls.gov; 202.691.7763)
- The Office of Administration is responsible for maintaining researcher' legal agreements and processing extensions.

## Data Access

### Onsite Appointments

Researchers should email their Project Coordinator at least two weeks in advance of a desired appointment in order to have a terminal reserved in the Researcher Suite. An appointment should be confirmed prior to making any travel arrangements. Starting October 15, 2017, badge access will no longer be granted to the Postal Square Building or the Researcher Suite. Your Project Coordinator will add your name to the Visitor Log System for the days on which you have a reservation. Please check in with the guards at the visitor entrance upon arrival at BLS.

If the Visitor Log System is down, a BLS staff member will need to meet you at the guard station and escort you upstairs. If your Project Coordinator is unavailable, the following individuals may be of assistance:

- John Jones (jones.john@bls.gov; 202.691.5197)
- Laurie Salmon ([salmon.laurie@bls.gov](mailto:salmon.laurie@bls.gov); 202.691.6511)

The Researcher Suite is available for use during regular business hours, Monday through Friday, from 8am to 5pm.

Note: If the Federal Government is closed (snow day, furlough, etc.), researchers are not permitted onsite, even if they have a standing appointment.

### Passwords

Each researcher has two sets of usernames and passwords: one for Windows, one for LINUX. A Windows password will remain valid for 90 days, and a LINUX password will remain valid for 84 days of activity or 60 days of inactivity. Researchers should be aware of these periods, and follow the procedures below if passwords have expired.

*Before coming to BLS:* Researchers must email their Project Coordinator - two days prior to their appointment in order to have passwords reset. If passwords have expired and are not reset, it may take several hours before access can resume.

*While at BLS:* If a researcher is "locked out" of an account, they should email or call their Project Coordinator as soon as possible. Do note that it may take several hours before new passwords can be generated.

## Research Server

The LINUX server for outside researchers is named CFResearch. In order to log into the server, launch Tectia-SSH Terminal and log into CFResearch using the username and password provided by your Project Coordinator. Note: upon your first log in you will receive a warning message, please simply click through.

## Research Server Uploads

All files are to be funneled through the OES Project Coordinator. No files may be uploaded directly by researchers. Files are to be emailed to the OES Project Coordinator prior to an onsite appointment. [Zip files may be sent by changing the .zip extension to a .zop extension. If you do not change the extension, files will be delayed and may be deleted by the BLS firewall.] If files are too large to email, they may be brought into BLS on portable media and uploaded by your Project Coordinator. Please be advised that this may take time and that it is always preferable to send files ahead of time. All files to be uploaded must be included in the scope of the Letter of Agreement or Memorandum of Understanding with the Researcher's institution.

## How to Transfer Files using Tectia Secure File Transfer

### Password Reset

- 1) **If you don't recall your password, ask your Project Coordinator to reset it prior to your arrival.** Open Tectia-SSH Terminal and connect to the server with the reset password provided to you by your Project Coordinator. You will be prompted to change your current password on the Tectia-SSH Terminal.

Click **File** -> **Quick Connect** -> Enter  
Host Name: CFResearch  
Port Number: 22

- Try the IP Address 10.112.31.3 if host name CFResearch isn't connecting.

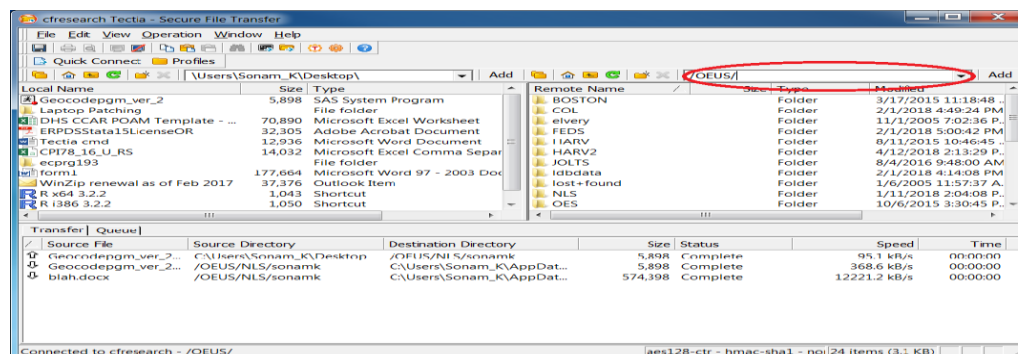
### Locate Your UNIX Folder

- 1) Open Tectia-Secure File Transfer (SFT) to locate your UNIX folder and connect to the server using your new password.

Click **File** -> **Quick Connect** -> Enter  
Host Name: CFResearch  
User Name: Contact your Project Coordinator  
Port Number: 22

- Try the IP Address 10.112.31.3 if host name CFResearch isn't connecting.

- 2) Enter “/OEUS” on your right search bar (circled below) and find your group. For example, an OES data user will click on “/OEUS/users/” and find their folder under their username. Select all your files and drag it to your left side of the Tectia. You may drag your files to “Desktop” or “Document.” Close Tectia and find your files on the workstation.



- If you don't see any files under your folder name, contact your Project Coordinator directly. Your Project Coordinator may need to make you the owner of your folder or your files may be hidden. To view hidden files on Tectia, Click **View > Show Hidden Files**.
- 3) At the end of your visiting day, you **MUST TRANSFER** all your files back to your UNIX folder and **DELETE** all those files from the workstation. Any files that are still on the workstation's hard drive will be **DELETED WITHOUT NOTICE**. To transfer files:
    - Connect CFRESEARCH on Tectia > Select your files on the workstation > Drag it to UNIX folder (/OEUS/group/yourusername) > Close Tectia. Depending on how big your files are, allow up to 30 minutes to complete these steps.
    - DELETE all the transferred files from the workstation -> Logoff.

### Running Programs Overnight

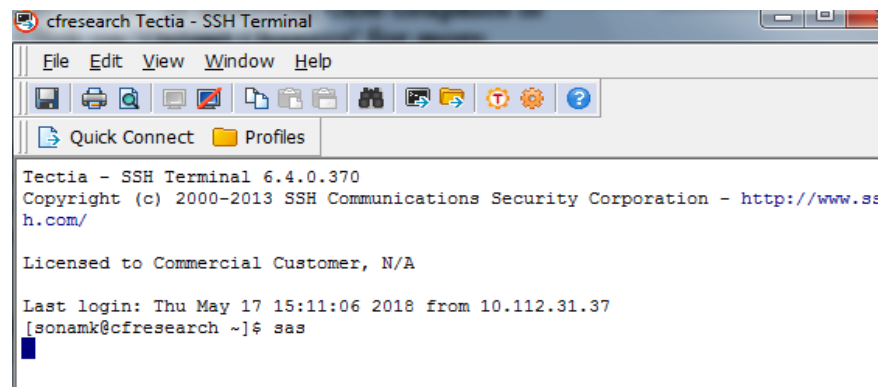
Let your Project Coordinator know that you will be running programs overnight so that no one uses the computer. Place a note next to your keyboard.

### Graphical User Interface (GUI) versions of SAS and STATA

Researchers may choose to use Graphical User Interface (GUI) versions of SAS and STATA on the research server. In order to enable the graphical display, please do the following:

How to use SAS and STATA on the CFResearch server

- 1) Enable X11 Tunneling
- 2) Start Tectia – SSH Terminal -> Click Edit -> Tectia Connections -> Default Connection -> Tunneling Tab -> Uncheck “Use factory defaults” -> Check “Tunnel X11 Connections” -> Click OK
- 3) Launch Exceed
  - a. Click on Window's Start -> Click All Programs -> Click on Open Text Exceed -> Open Exceed
- 4) Connect to CFResearch server on Tectia-SSH Terminal
- 5) At the CFResearch prompt, enter sas. The SAS application will appear in the Exceed window:

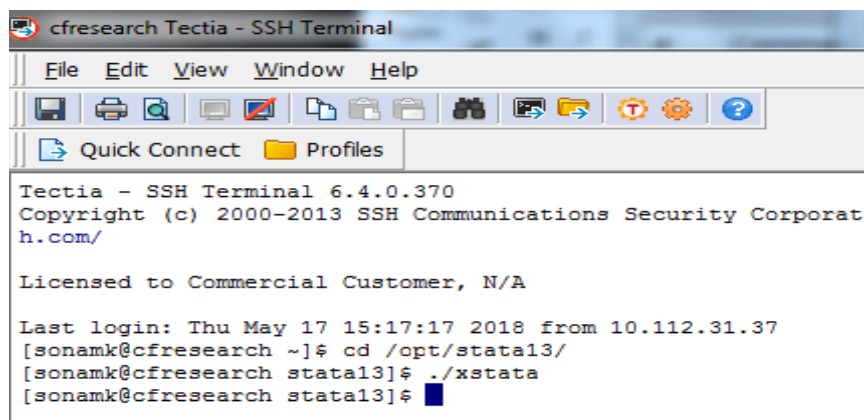


```
Tectia - SSH Terminal 6.4.0.370
Copyright (c) 2000-2013 SSH Communications Security Corporation - http://www.ssh.com/

Licensed to Commercial Customer, N/A

Last login: Thu May 17 15:11:06 2018 from 10.112.31.37
[sonamk@cfresearch ~]$ sas
```

- To use STATA 12 -> Enter xstata on the CFResearch prompt. The default STATA 12 will open.
- To use STATA 13 -> Enter cd/opt/stata13/ -> ./xstata



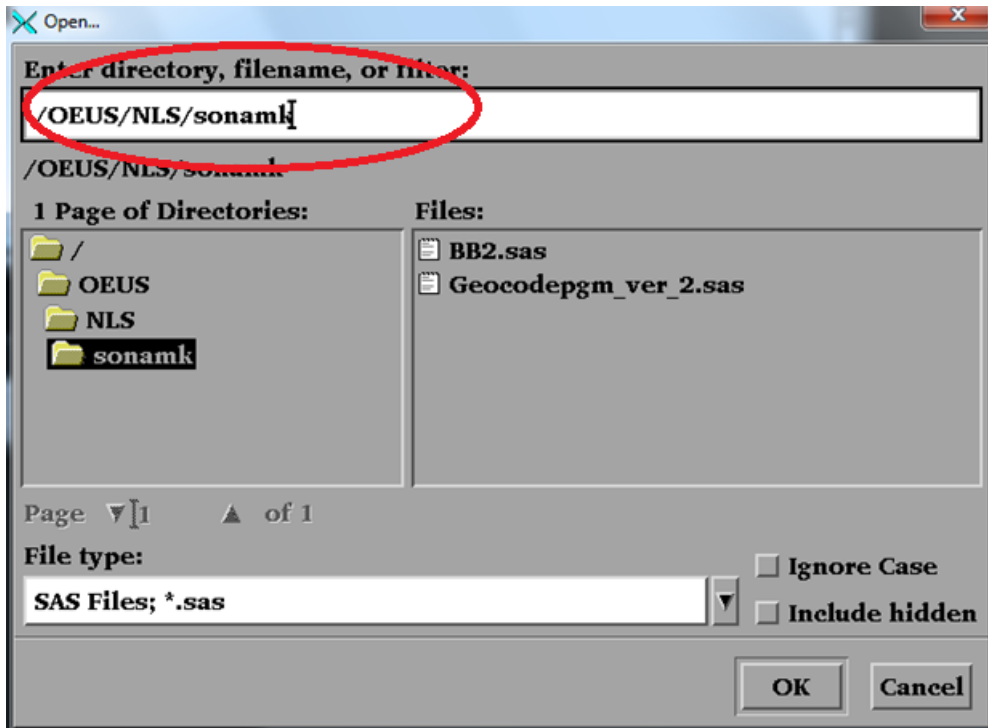
```
Tectia - SSH Terminal 6.4.0.370
Copyright (c) 2000-2013 SSH Communications Security Corporation - http://www.ssh.com/

Licensed to Commercial Customer, N/A

Last login: Thu May 17 15:17:17 2018 from 10.112.31.37
[sonamk@cfresearch ~]$ cd /opt/stata13/
[sonamk@cfresearch stata13]$ ./xstata
[sonamk@cfresearch stata13]$
```

#### Navigating Graphical User Interface (GUI) of SAS

1. Opening SAS files: On SAS editor window, click File -> Open -> Enter file directory (/OEUS/GROUP/Username) -> Find your files



2. Saving SAS files: If you are saving the log, output, or editor windows -> File -> Save As -> Make sure you are saving it on the /OEUS/group/username -> name your file -> Check to see if the files are there on Tectia - Secure File Transfer

#### STATA and R packages

- 1) Click Start -> Computer -> "C Drive" -> Open "STATA Packages" or "R Packages" folders to find the packages.
- 2) If your packages aren't listed in the folders, contact your Project Coordinator directly. Your Project Coordinator should submit a request to install the packages.
- 3) Researchers are prohibited from uploading any packages or files on their USB drives.

## Information for Outside Researchers Using OES Microdata

### Location of the files

The available OES files are located on the RESEARCH UNIX server in the folder /OEUS/users/OES\_data/. All of the data files are zipped, so they must be unzipped before use. The /OES\_data folder is read-only, so the files must be moved to another workspace to use them.

### Types of files available

**Mastdata files:** These are the main data files that are used to create the OES estimates. The files are in SAS dataset format. The files contain all of the occupational employment and wage data, along with the weighting and other information needed to produce the OES estimates. The wage data in the Mastdata files are in wage intervals. The Yrints files are used to convert the wage interval values into actual wage rate values.

The Mastdata files each contain 6 panels or 3 years of data that was used to produce the OES estimates referenced in the Mastdata file name. For example, the 201402 Mastdata file includes survey data for 201104, 201202, 201204, 201302, 201304, and 201402. This 201402 Mastdata was used to produce the May 2014 OES estimates. The 201302 Mastdata file contains survey data from 201004-201302. It is possible that some of overlapping data in two or more Mastdata files could be different due to data corrections or some data being dropped due to re-sampled units. For example, the 201204 panel data in the 201402 Mastdata file may be different than the 201204 panel data in the 201302 Mastdata due to some additional corrections that came in after the May 2013 estimates were produced.

**Yrints files:** These files contain the information needed to convert the interval wage data to actual wage rates. Each Mastdata file has a corresponding Yrints table as shown by the year in the file name. For example, the Yrints\_201402 dataset corresponds to the Mastdata201402 dataset. The yrints file should be merged or linked to the mastdata file by the fields Yr+Fips+occ\_code. Once linked, the IntMu\_1 value from yrints shows the interval mean for Wage A in the mastdata file; IntMu\_2 shows interval mean for Wage B, ... IntMu\_12 is the interval mean for Wage L. All of the interval means are aged so that they represent the panel of the estimates even for the older data in the mastdata dataset.

**E0 files:** These files contain additional establishment information that might not be contained in the mastdata files, such as establishment names. Each E0 file only contains data for the year referenced in the file name. For example, the E0\_20142 dataset only contains 201402 schedules. To have E0 files for all of the schedules in the 201402 Mastdata file, the six E0 files for 20114 – 20142 need to be used. The E0 files should be merged with the Mastdata files using the Schnum field. The Schnum is distinct for each establishment in a given mastdata file.

### Reference files

Reference files that provide field description information and other information about the data files are available on the CFRESEARCH server in the folder /OEUS/users/OES\_data/Reference\_files/

## Non-Disclosure Review

One or more people will be assigned to review and clear your statistical work and related program files before it is released to you. We review it for disclosure risk and statistical work that is not related to the project proposal. You provide the filenames for the output files. The reviewer may do some confidentiality checks or redaction of output.

The only things we can usually release are:

- 1) Statistical analyses – regression or other type of estimation
- 2) Data summaries – tables of means for example
- 3) Programs – without data

This means that no microdata can be taken from the Researcher Suite, no matter how well masked.

The following points may be relevant to review of output on your project:

1. Review takes up to two weeks under normal circumstances. Large volumes take longer.
2. Difficult review questions are directed to an internal review board.
3. We cannot divulge our exact confidentiality rules, lest someone reverse engineer a confidential piece of data. We can describe the goal of masking information that would otherwise identify individual responses.
4. Revealing that the sample contains members who reside in a particular geographic unit with population less than 5% of the total US population is a confidentiality breach. For example, researcher output or programs cannot be taken out if they include information that indicates the number of observations by county, zip code, or census tract. Information on sample sizes at the state or state aggregate level can be taken out only if the current population of the state or state aggregate exceeds 5% of the total US population.
5. You cannot reveal individual observations' microdata values. For example, this means that you should redact the minimum and maximum values typically produced with summary statistics (such as the SAS MEANS procedure or the Stata summarize command). We also encourage you to redact ancillary or non-critical statistics, for example regression coefficients associated with background control variables.
6. We occasionally have to actually make computations to adequately review output. Therefore we need access to and a basic understanding of your programs and data.
7. Your agreement may include approval to match BLS and external data. You should not take cell-average data (or regression-prediction data) out in order to match to external data files. You should instead bring any external data to BLS, match here, and calculate the ultimate statistics of interest for us to review.
8. We may have to limit release of very similar statistical runs. You should certainly undertake robustness checks, but should try to limit the amount of similar output that must be reviewed.
9. We must limit release of tabular output. Tables with thin cells are difficult and time consuming to review for confidentiality. Furthermore, researchers typically use these tables only as intermediate output.
10. Output should be restricted to the approved project and no other topics.
11. If you believe that these rules have been arbitrarily applied by your reviewer, you can discuss your case with the OEUS representative to the BLS Microdata Access Review Board.



Please try to limit the intermediate output you take out. Output review can be very burdensome to us in terms of the time it requires and constitutes a major cost to us in running the Visiting Researchers program. With a little bit of care and a willingness to do a little more of their work onsite, researchers can generally sharply limit the intermediate output that needs to be reviewed. Reducing the volume of material to be reviewed will allow us to review your review much more quickly.

Please provide your Project Coordinator copies of all papers that use BLS microdata; these provide a measure of the benefits of the researcher access program.

Researchers are required to include the following disclaimer in any published works using these data: *This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed here do not necessarily reflect the views of the BLS or the U.S. government.*

### [Project Amendments](#)

#### [Project Extensions](#)

Project extensions are granted on a case by case basis. Researchers may begin the extension application process 3-6 months prior to their agreement's expiration. Justification for the extension, as well as a summary of work to be completed will be required. If Researchers have completed their work, but are in the process of finalizing their publications, they may enter a referee comments period for 1 year after their agreement has expired. During this year, a researcher may then apply for a 3 month timeframe to make minor revisions to their outputs.

Note: Arrangement for a project extensions and referee comments period must be made through the Office of Administration prior to the project's expiration.

#### [Transferring Sponsoring Institutions](#)

Researchers may transfer institutions, but some restrictions apply. Please refer to the Institution Transfer Policy for details.

## Safety

### Building Evacuation

In the event of a building evacuation, researchers are to quickly and safely leave the building, and join the employees of Room 4945 in Capitol Hill Park.

- Exit Room 4955 and head towards Stairway #5, located 108ft to the right of the office. Follow the directions of emergency monitors, who will be wearing red baseball caps that say Postal Square OEP Monitor on the brim.
- Exit the building from the first floor at the corner of North Capitol Street & Massachusetts Avenue. From the ground floor, exit onto North Capitol Street (follow sign). Exit through the stairway by suite 1950.
- Follow the emergency monitors to Capitol Hill Park which is located 620ft from the building at the corner of Massachusetts and Louisiana Avenues, N.E.

### Shelter In Place

During a shelter in place or a shelter in place drill, researchers in room 4955 are to go to room 4930. Exit the room 4955 door and room 4930 is across the hall and to the right. Room 4930 is marked with a square yellow sign that reads: Shelter in Place.