
AWS - RStudio Server Installation Instructions

*Daniel Hammocks **
UCL Department of Security and Crime Science
2018-10-03
Version 2

Abstract

This document covers the use of Amazon Web Services (AWS), a scalable cloud computing service, for compiling code written in R by utilising a Linux based RStudio server. The instructions primarily focus on the following three topics; AWS account setup, RStudio server installation and RStudio GUI usage. This is a comprehensive guide, complete with screenshots, so the length of this document is misleading to the complexity of the entire process which can be completed in less than twenty minutes.

Contents

1	Introduction	3
1.1	What is Amazon Web Services?	3
1.2	What is RStudio Server?	3
2	AWS Account Creation	4
3	AWS Account Login	7
4	AWS RStudio Server Setup	8
5	Accessing AWS Instance Via SSH	14
5.1	Installing PuTTY	14
5.2	PuTTYgen Key Conversion	15
5.3	Connecting via SSH	17
5.4	Changing RStudio GUI User Password	18
5.5	Adding Users	19
6	Using RStudio Server Graphical User Interface	21

* daniel.hammocks.18@ucl.ac.uk

7 Uploading Files via FTP	23
7.1 Installing FileZilla Client	23
7.2 Changing Folder Permissions Via SSH	24
7.3 Connecting to AWS Instance Via FTP	25
7.4 Adding Files for Use in R	27

1 Introduction

1.1 What is Amazon Web Services?

Overview Arguably, one of the most important aspects of data science is the speed at which data can be analysed. There are a number of limiting factors that can affect the computational speed including the way in which code is written, the programming language used and the hardware resources available.

The purpose of Amazon Web Services (AWS), and cloud computing in general, is to remove the latter limitation by replacing the expensive hardware ownership cost, associated with access to large scale computing resources, with a significantly cheaper pay-as-you-compute pricing model.

Usage and Benefits AWS identify six main benefits of using their services over traditional compute servers. These can be summarised as:

- Lower Costs
- Scalable Platforms
- Increased Speed
- No Hardware Maintenance
- Simple to Use and Implement

Limitations Whilst there are many appealing characteristics of cloud based computing you should also consider; the security and privacy of your data including the ownership of data that is uploaded to a cloud service, the vulnerability of running online computations, and the increasing costs as your usage increases.

1.2 What is RStudio Server?

What is RStudio? RStudio is an open source desktop application that makes the programming language R easier to use by providing a user friendly graphical user interface (GUI), debugging tools, visualisation tools, and a code editor.

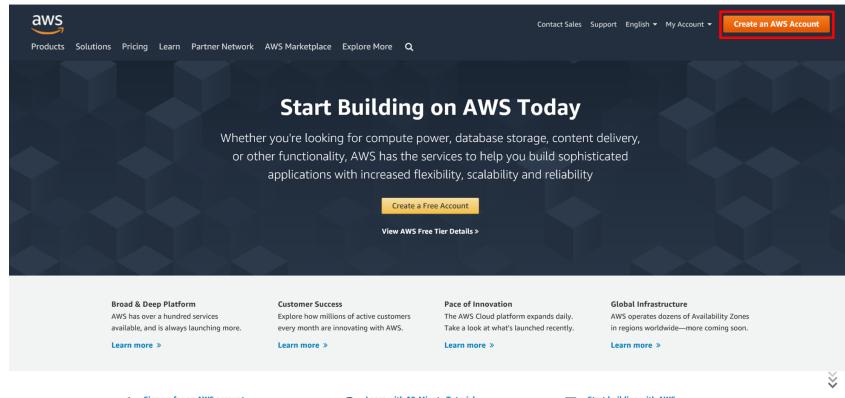
Put simply, RStudio is a program that enables you to edit and compile code written in R.

Overview RStudio Server is an open source (Linux server application) variant of the RStudio desktop version. The server edition can be installed on a server, with high compute capabilities, and accessed via a web browser.

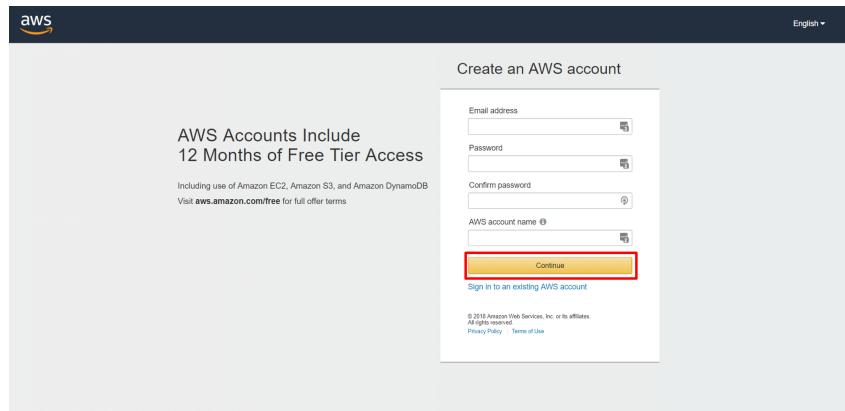
2 AWS Account Creation

1. Navigate to <https://aws.amazon.com>

2. Click on 'Create an AWS Account'.



3. Complete the initial sign-up information and click 'Continue'.



4. You will be redirected to a page requesting additional contact information. Complete these fields.

5. Before you continue I advise you read the AWS User Agreement (<https://aws.amazon.com/agreement/>) and the AWS Data Privacy Facts and Questions (<https://aws.amazon.com/compliance/data-privacy-faq/>).

6. Once you have ensured AWS is suitable for your usage purpose click 'Create Account and Continue'.

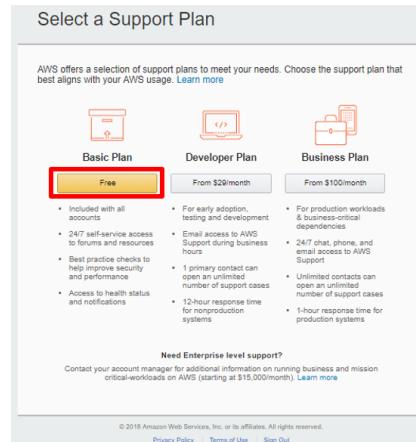
7. You will be redirected to a page requesting payment information. Please note that you will only be charged if you exceed the free usage limit. Complete these details and then click 'Secure Submit'.

8. You will now be required to verify your contact number. Enter the characters shown as part of the bot security check and click 'Call Me Now'.

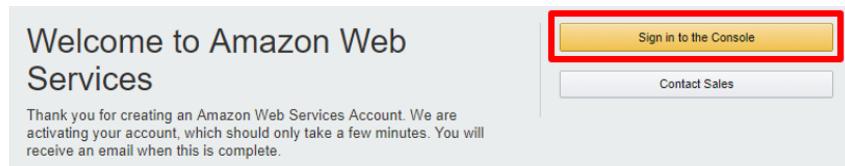
9. You should now receive a call from AWS. Answer the call and enter, using your phones keypad, the four digits shown on your screen.



- When you have successfully verified your contact details you will need to select a support plan. Choose the Basic Plan by clicking 'Free'.

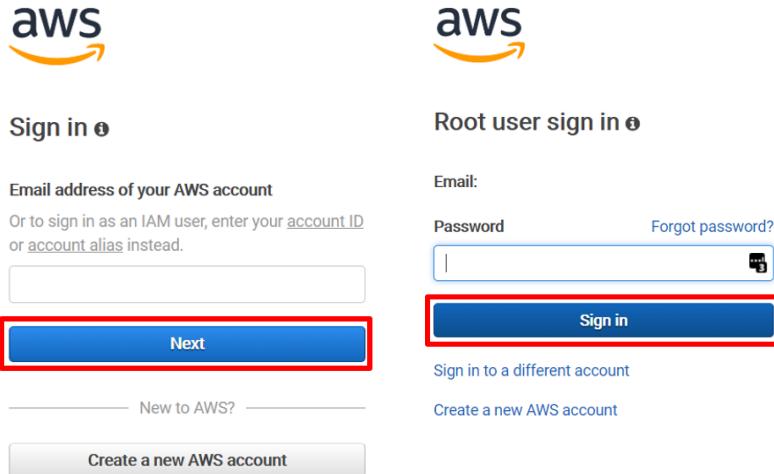


- You have now successfully setup your AWS account. Click on 'Sign in to the Console' to proceed to the AWS setup.

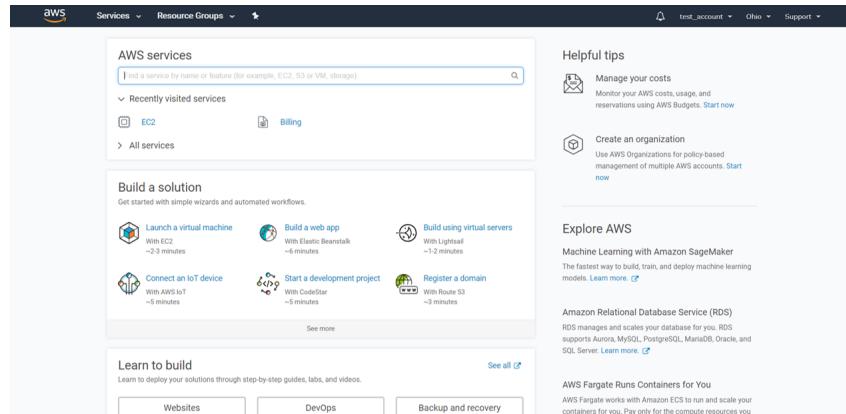


3 AWS Account Login

1. In order to login to the AWS console you must visit <https://aws.amazon.com/console/>.
2. Enter your email address. Click 'Next'. Then enter your password followed by clicking the 'Sign in' button.



3. You should then reach the following page.



4 AWS RStudio Server Setup

1. Access the AWS Console.

The screenshot shows the AWS Home Page. At the top, there's a search bar with placeholder text "Find a service or feature (example, EC2, S3 or VM, storage)". Below the search bar, there are sections for "Recently visited services" (EC2, Billing), "Build a solution" (Launch a virtual machine, Build a web app, Build using virtual servers, Connect an IoT device, Start a development project, Register a domain), and "Learn to build" (Websites, DevOps, Backup and recovery). On the right side, there are "Helpful tips" (Manage your costs, Create an organization), "Explore AWS" (Machine Learning with Amazon SageMaker, Amazon Relational Database Service (RDS), AWS Fargate Runs Containers for You), and other navigation links like "test_account", "Ohio", and "Support".

2. In the services search bar type 'EC2' and then click on the respective option. This will take you to the EC2 dashboard.

The screenshot shows the AWS Services search results. A red box highlights the search bar at the top with the text "ec2". Below it, the "EC2" service is listed under "Virtual Servers in the Cloud". Other services listed include EFS, Elastic Container Service, and GuardDuty. At the bottom, there are three buttons: "Launch a virtual machine", "Build a web app", and "Build using virtual servers".

3. On the EC2 dashboard click 'Launch Instance'.

The screenshot shows the EC2 Dashboard. On the left, there's a sidebar with links for EC2 Dashboard, Events, Tags, Reports, Limits, Instances, Images, Elastic Block Store, Network & Security, and Key Pairs. The main area shows "Resources" for the US East (Ohio) region, including 0 Running Instances, 0 Dedicated Hosts, 0 Volumes, 0 Key Pairs, and 0 Placement Groups. To the right, there are sections for "Account Attributes" (Support Platforms, VPC, Default VPC), "Additional Information" (Getting Started Guide, Documentation, All EC2 Resources, Forums, Pricing, Contact Us), and "AWS Marketplace" (Find free software trial products in the AWS Marketplace from the EC2 Launch Wizard, or try these popular AMIs). A central box says "Create Instance" and "To start using Amazon EC2 you will want to launch a virtual server, known as an Amazon EC2 instance." It features a "Launch Instance" button, which is highlighted with a red box. Below it, there's a note: "Note: Your instances will launch in the US East (Ohio) region".

4. We will use an Amazon Machine Image (AMI) created and maintained by Louis Aslett (<http://www.louisaslett.com/>). Type 'rstudio' into the search bar. Click on the 'Community AMIs' option.

- Then select the 'RStudio + Julia/Python server' variant.

- You will then be asked to define the instance type which determines the number of CPUs, RAM and storage available to the instance as well as the network performance. For the purpose of this tutorial we will select the 'General purpose Free Tier Eligible' option.

- Click 'Next: Configure Instance Details'.

- We will not alter any of the instance settings in this instruction manual. Click 'Next: Add Storage'.

9. On this page you can select the size and type of storage available to the instance. Please note that the free account is restricted to, at the time of writing, 30Gb of SSD storage.

10. Click 'Next: Add Tags'

11. Click 'Next: Configure Security Group'

12. We will now 'Create a new security group'. These settings determine which traffic is allowed to access the instance. Since we want to be able to access our server from locations with Dynamic IP addresses we will remove HTTP and HTTPS restrictions.

(a) Select 'Create a new security group'.

- (b) Define the 'Security group name' as HTTP/S Anywhere.
- (c) Use the following description so that the security group can be reused if deemed suitable. 'Access Instance Via SSH/HTTP/HTTPS - No IP address restriction'

- (d) Click 'Add Rule'
- (e) Under the new rule, select type 'HTTP'.

- (f) Change the source to 'Anywhere'.

- (g) Click 'Add Rule'
- (h) Under the new rule, select type 'HTTPS' and change the source to 'Anywhere'.

13. Click 'Review and Launch' to proceed.

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more about Amazon EC2 security groups.](#)

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	Custom • 0.0.0.0/0	e.g. SSH for Admin Desktop
HTTP	TCP	80	Anywhere • 0.0.0.0/0	e.g. SSH for Admin Desktop
HTTPS	TCP	443	Anywhere • 0.0.0.0/0	e.g. SSH for Admin Desktop

Add Rule

Warning:
Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

Review and Launch

14. Review the settings, if required, and then click 'Launch'.

Step 7: Review Instance Launch

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

AMI Details

RStudio-1.1.383_R-3.4.2_Julia-0.6.0_CUDA-8_cuDNN-6_ubuntu-16.04-LTS-64bit - ami-3b0c205e
Ready to run RStudio + Julia/Python server for statistical computation ([www.louisaslett.com](#)). Connect to instance public DNS in web browser (standard port 80), username rstudio and password rstudio
Root Device Type: ebs Virtualization type: hvm

Instance Type

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
t2.micro	Variable	1	1	EBS only	-	Low to Moderate

Security Groups

Launch

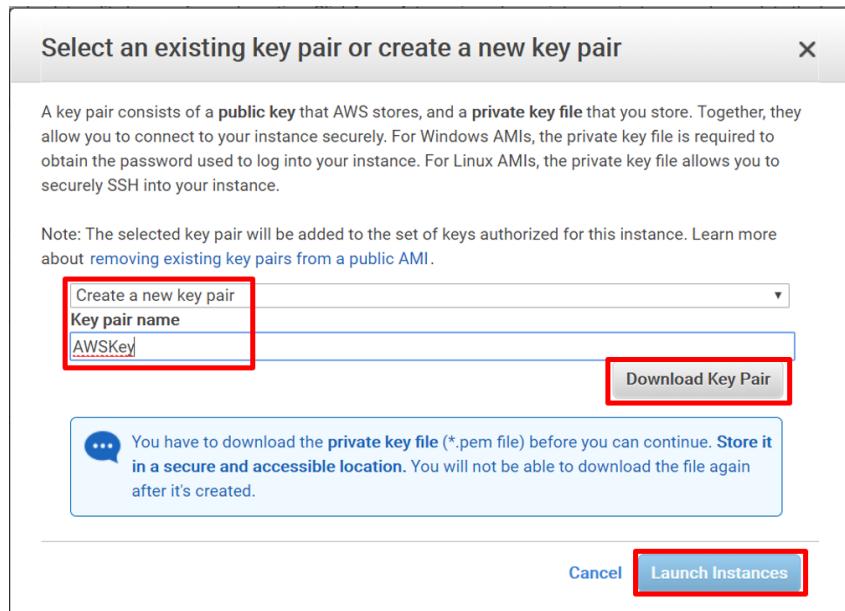
15. You will then be asked to select or generate a key pair.

(a) Select 'Create a new key pair'.

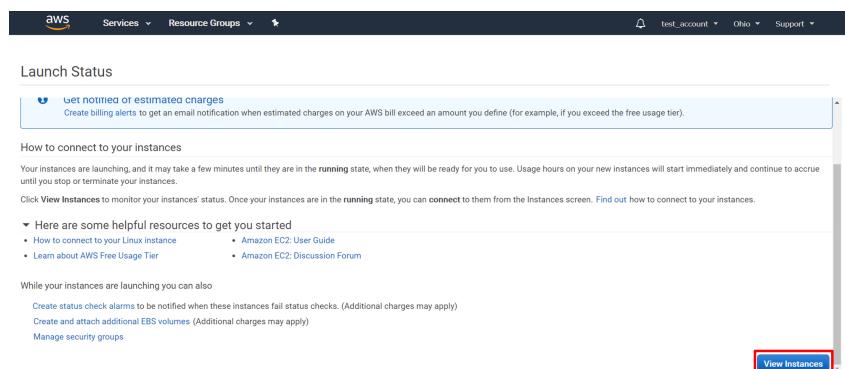
(b) Define a name for the new key.

(c) Then click 'Download Key Pair'. You will need this later to access the instance via SSH - keep it safe.

16. Click 'Launch Instances'.



17. Scroll down and click 'View Instances'.



5 Accessing AWS Instance Via SSH

5.1 Installing PuTTY

In order to access the AWS instance via SSH you will need an SSH client such as PuTTY. The following instructions are for Windows users only. A similar process can be completed on Linux devices. Mac users will need to use the Mac OS X built-in SSH client.

1. Navigate your web browser to <https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html> and download the appropriate version of PuTTY for your computer.

Download PuTTY: latest release (0.70)

Home | FAQ | Feedback | Licence | Updates | Mirrors | Keys | Links | Team
Download: Stable Snapshot | Docs | Changes | Wishlist

This page contains download links for the latest released version of PuTTY. Currently this is 0.70, released on 2017-07-08.
When new releases come out, this page will update to contain the latest, so this is a good page to bookmark or link to. Alternatively, here is a [permanent link to the 0.70 release](#).
Release versions of PuTTY are versions we think are reasonably likely to work well. However, they are often not the most up-to-date version of the code available. If you have a problem with this release, then it might be worth trying out the [development snapshots](#), to see if the problem has already been fixed in those versions.

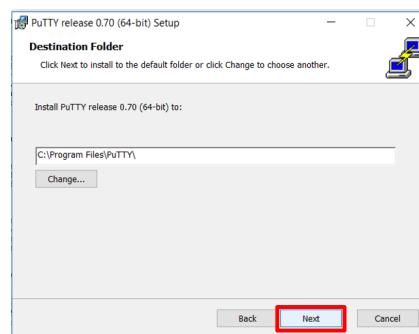
Package files		
You probably want one of these. They include all the PuTTY utilities. (Not sure whether you want the 32-bit or the 64-bit version? Read the FAQ entry .)		
MSI (Windows Installer)		
32-bit: putty-0.70-installer.exe	(or by FTP)	(or git (tarball))
64-bit: putty-0.70-1a-0.70-installer.exe	(or by FTP)	(or git (tarball))
Unix source archive		
putty-0.70-tar.gz	(or by FTP)	(or git (tarball))

Alternative binary files		
The installer packages above will provide all of these (except PuTTYtel), but you can download them one by one if you prefer.		

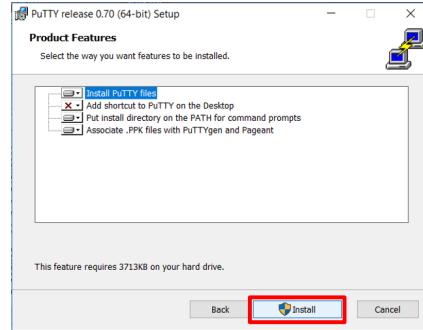
2. Open the installation executable and run through the setup process.
 - (a) Click 'Next'.



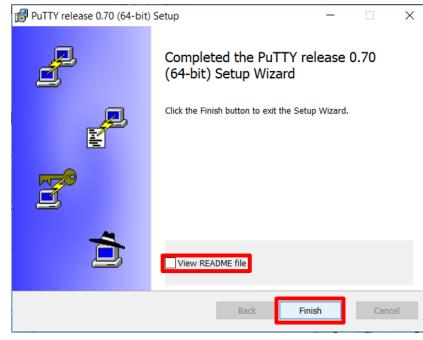
- (b) Click 'Next'.



- (c) Click 'Install' (You may need to provide access to the program through Windows UAC)



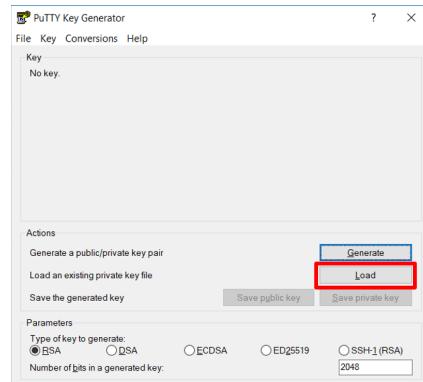
- (d) Untick 'View README file' and then click 'Finish'.



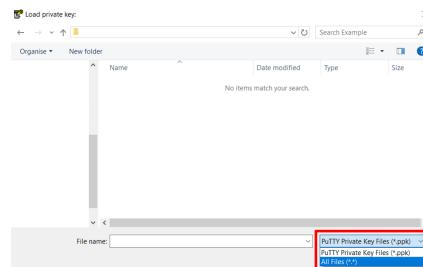
5.2 PuTTYgen Key Conversion

Before you can connect to the instance via SSH you need to convert the key you downloaded into a format that can be understood by PuTTY.

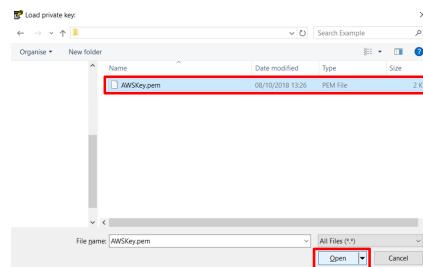
1. Open PuTTYgen - a program that is installed alongside PuTTY.
2. Click 'Load'.



3. Ensure that the valid file types is set to 'All Files (*.*)'.

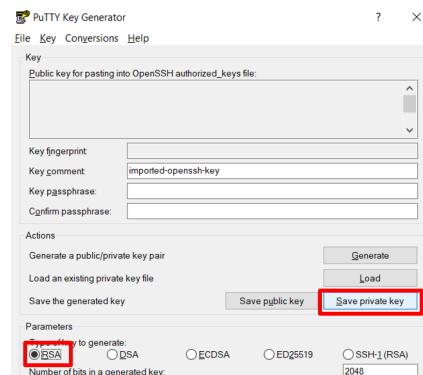


4. Select the key that you saved from AWS. Then click 'OK'.

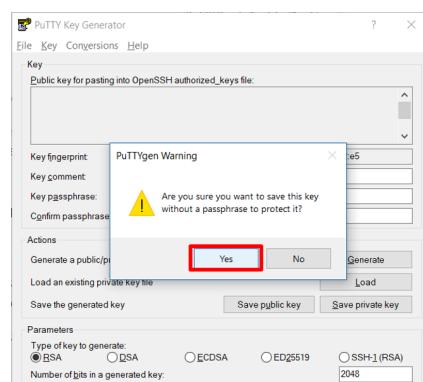


5. Under parameters set the type of key to generate as 'RSA'.

6. Then click save private key.



7. Click 'Yes'.

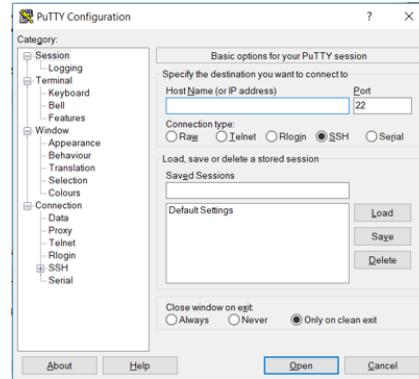


8. Save the new key with the same name as the old key.

9. You can now close the PuTTYgen program.

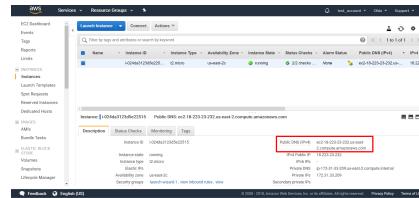
5.3 Connecting via SSH

1. Open PuTTY.



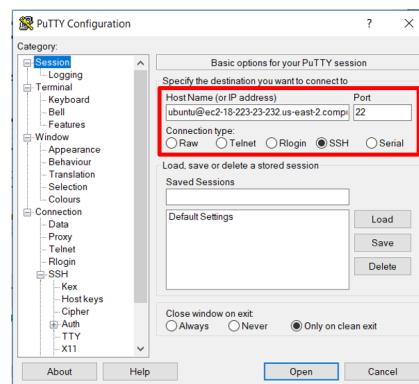
2. Enter the following details

- (a) Host Name: 'ubuntu@###' where ### is the Public DNS (IPv4) from the EC2 Instance Dashboard shown below.



- (b) Port: 22

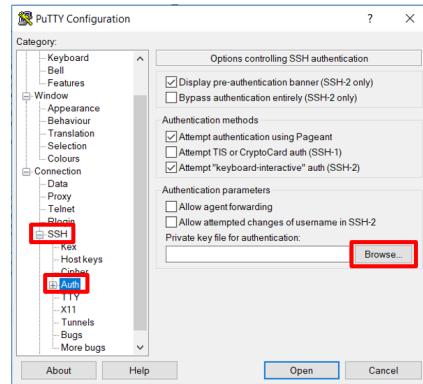
- (c) Connection Type: SSH



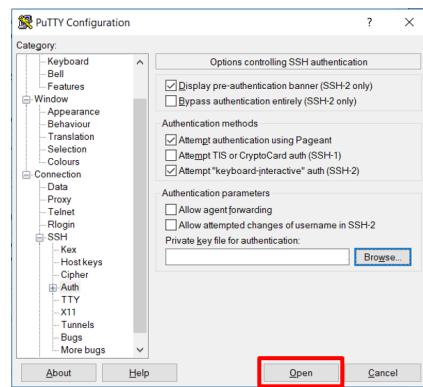
3. On the left hand side click on 'SSH'.

4. Under 'SSH' click on 'Auth'.

5. Then click 'Browse' and locate the key file we created above.

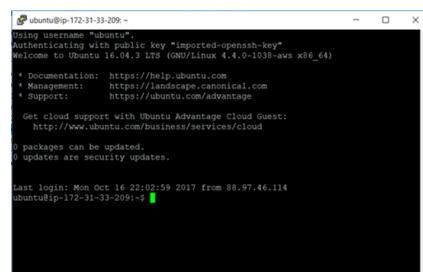


6. Then click 'Open'.



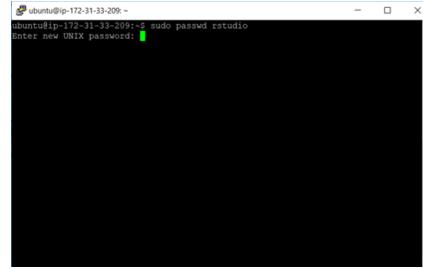
7. A warning dialogue box will appear. Click 'Yes'.

8. You should then get a command line window as shown below.

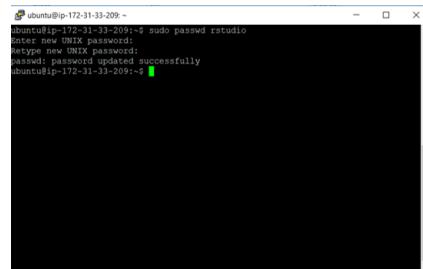


5.4 Changing RStudio GUI User Password

1. Access the Instance via SSH as above.
2. Enter 'sudo passwd username' where username is the user name of the account you wish to change the password of. The default account is 'rstudio'. The password for the default account should be changed during setup.

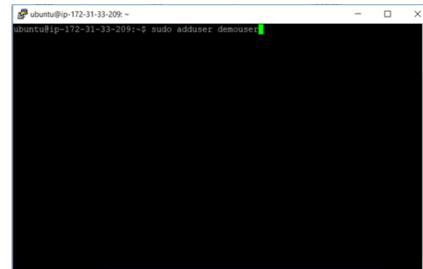


3. You will then be required to input the desired password twice. Type the desired password followed by the enter key. Please note that all keystrokes are registered but not displayed.

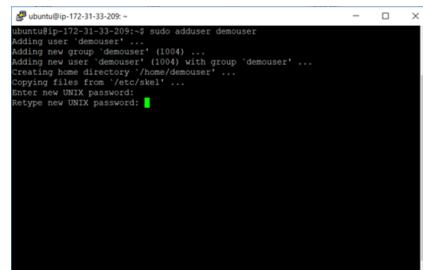


5.5 Adding Users

1. Access the Instance via SSH as above.
2. Enter 'sudo adduser username' where username is the user name of the account you wish to add.



3. You will then be required to input the desired password twice. Type the desired password followed by the enter key. Please note that all keystrokes are registered but not displayed.



4. Press the enter key five times.



```
ubuntu@ip-172-31-33-209:~$ sudo adduser demouser
Adding user 'demouser' ...
Adding new group 'demouser' (1004) ...
Adding new user 'demouser' (1004) with group 'demouser' ...
Creating home directory '/home/demouser' ...
Copying files from '/etc/skel' ...
Enter new UNIX password:
Retype new UNIX password:
password: password updated successfully
Changing the user information for demouser
Enter the new value, or press ENTER for the default
    Full Name []:
    Room Number []:
    Work Phone []:
    Home Phone []:
    Other []
Is the information correct? [Y/n] n
```

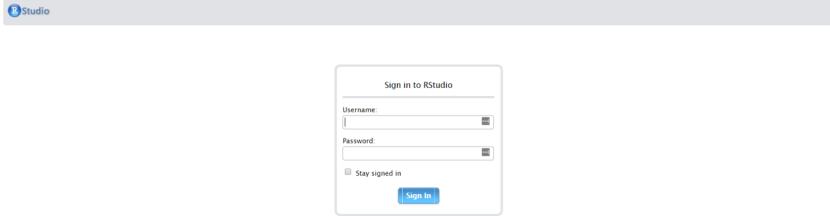
5. Then type 'y' followed by the enter key.



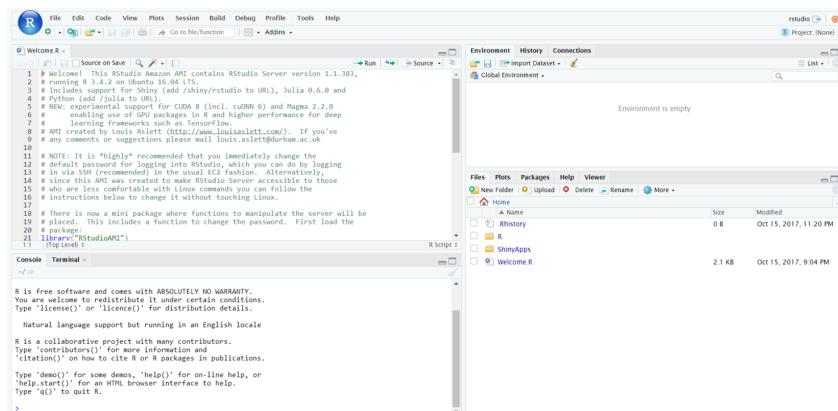
```
ubuntu@ip-172-31-33-209:~$ sudo adduser demouser
Adding user 'demouser' ...
Adding new group 'demouser' (1004) ...
Adding new user 'demouser' (1004) with group 'demouser' ...
Creating home directory '/home/demouser' ...
Copying files from '/etc/skel' ...
Enter new UNIX password:
Retype new UNIX password:
password: password updated successfully
Changing the user information for demouser
Enter the new value, or press ENTER for the default
    Full Name []:
    Room Number []:
    Work Phone []:
    Home Phone []:
    Other []
Is the information correct? [Y/n] y
ubuntu@ip-172-31-33-209:~$
```

6 Using RStudio Server Graphical User Interface

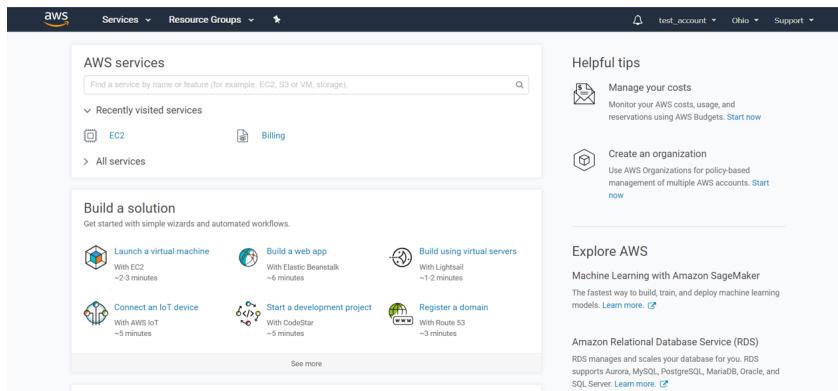
1. In your web browser navigate to the 'Public DNS (IPv4)' address as used above.
2. You should see the following page.



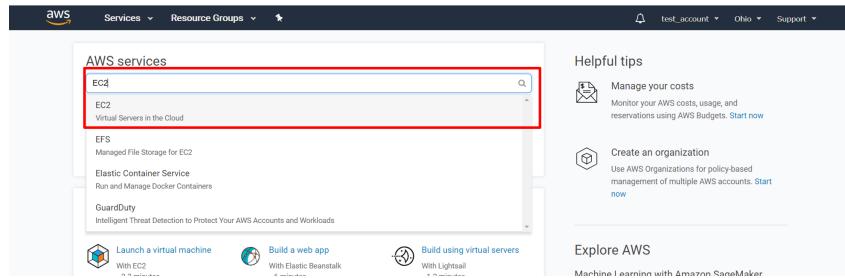
3. You can login using:
 - (a) Username: rstudio
 - (b) Password: 'Created Above in SSH Client by Yourself'
4. Once you have logged in you will see a GUI interface similar to that of the RStudio Desktop product.



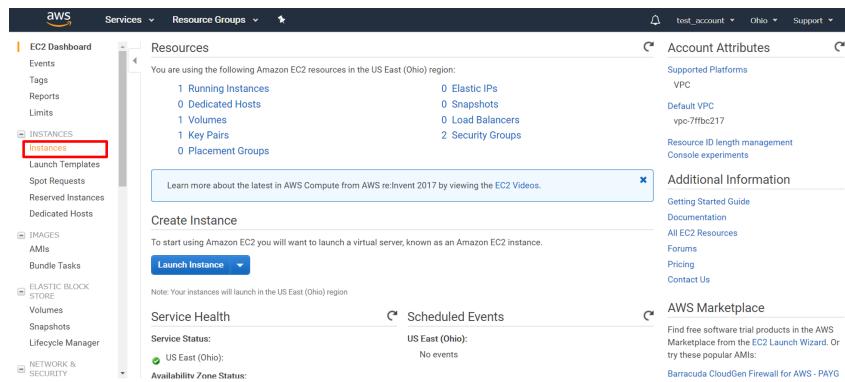
5. Once you have finished using the RStudio Server you should:
 - (a) Navigate to the AWS Console (<https://console.aws.amazon.com>).



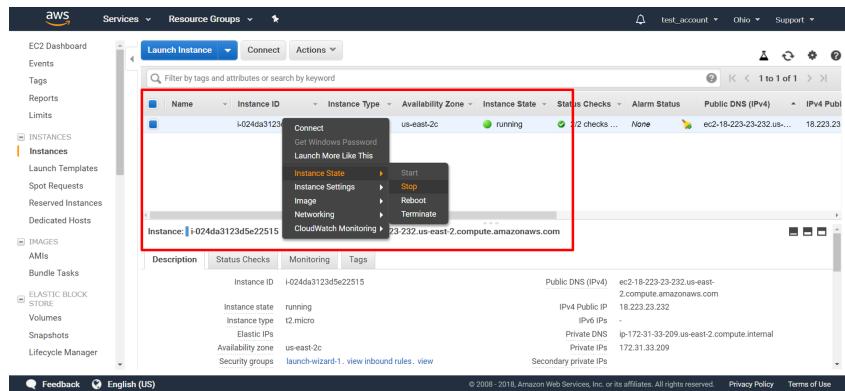
(b) Go to the EC2 dashboard.



(c) Click on 'Instances' on the left hand side.



(d) Then, right click on the instance and select 'Instance State' = 'Stop'.



This will prevent usage of your free computation hours when you're not directly using the server instance.

6. You can restart the service by repeating steps (a) to (d) above but by choosing 'Start' rather than 'Stop' in step (d).

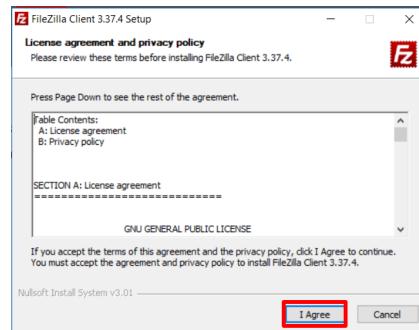
7 Uploading Files via FTP

7.1 Installing FileZilla Client

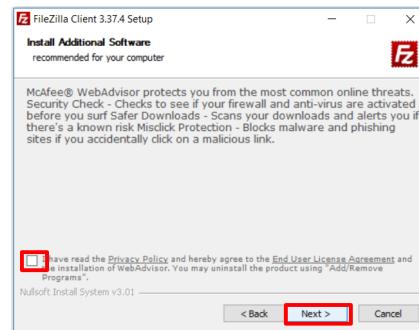
1. Navigate to <https://filezilla-project.org/download.php?type=client> to download the client.



2. Open the downloaded executable file.
3. Run through the installation process.
 - (a) Click 'I Agree'.



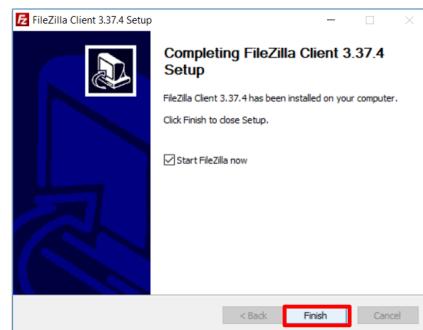
- (b) Click 'Next' four times.
- (c) Uncheck the additional software installation. Then click 'Next'.



- (d) Uncheck the additional software installation. Then click 'Next'.



(e) Click 'Finish'.



7.2 Changing Folder Permissions Via SSH

1. Connect to the instance via SSH as described above.
2. Use the 'cd' command to change directory to the home directory: 'cd /home'.

```
ubuntu@ip-172-31-33-209: ~
ubuntu@ip-172-31-33-209:~$ cd /home
```

3. You can use the 'dir' command to obtain a list of the user home folders.

```
ubuntu@ip-172-31-33-209: /home
ubuntu@ip-172-31-33-209:~$ cd /home
ubuntu@ip-172-31-33-209:/home$ dir
demouser rstudio shiny test testuser ubuntu
ubuntu@ip-172-31-33-209:/home$
```

4. Enter the following command to change the folder permissions: 'sudo chmod 777 -R /home/username' where username is the name of the user folder you wish to adjust permissions for - most likely 'rstudio'.

```

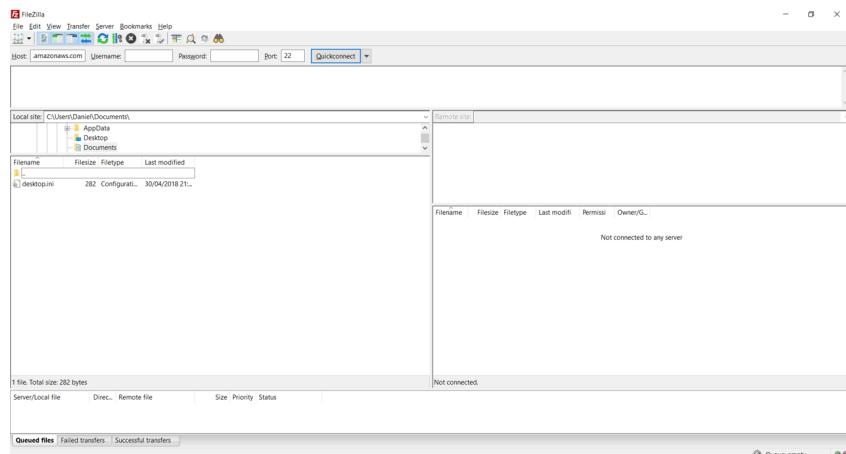
ubuntu@ip-172-31-33-209:~$ cd /home
ubuntu@ip-172-31-33-209:/home$ dir
demo user rstudio shiny test testuser ubuntu
ubuntu@ip-172-31-33-209:/home$ ls -la /home/rstudio
total 48
drwxrwxrwx 8 rstudio rstudio 4096 Oct 16 2017 .
drwxr-xr-x 8 root root 4096 Oct 8 15:10 ..
-rwxrwxrwx 1 rstudio rstudio 220 Aug 31 2015 .bash_logout
-rwxrwxrwx 1 rstudio rstudio 3771 Aug 31 2015 .bashrc
drwxrwxrwx 2 rstudio rstudio 4096 Oct 15 2017 .dropbox-ami
drwxrwxrwx 3 rstudio rstudio 4096 Aug 22 2017 .dropbox-disk
-rwxrwxrwx 1 rstudio rstudio 655 May 16 2017 .profile
drwxrwxrwx 3 rstudio rstudio 4096 Oct 9 08:29 .
-rwxrwxrwx 1 rstudio rstudio 0 Oct 8 18:46 .Rhistory
drwxrwxrwx 14 rstudio rstudio 4096 Oct 8 18:46 .rstudio
drwxrwxrwx 4 rstudio rstudio 4096 Oct 8 15:21 ShinyApps
drwxrwxrwx 3 rstudio rstudio 4096 Oct 15 2017 .subversion
-rwxrwxrwx 1 rstudio rstudio 2141 Oct 15 2017 Welcome.R
ubuntu@ip-172-31-33-209:/home$ sudo chmod 777 -R /home/rstudio
ubuntu@ip-172-31-33-209:/home$ 

```

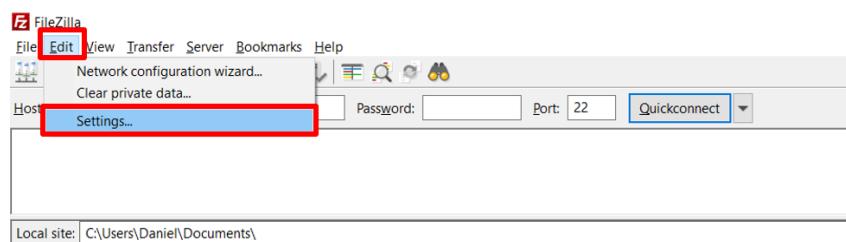
5. You can now close the SSH client.

7.3 Connecting to AWS Instance Via FTP

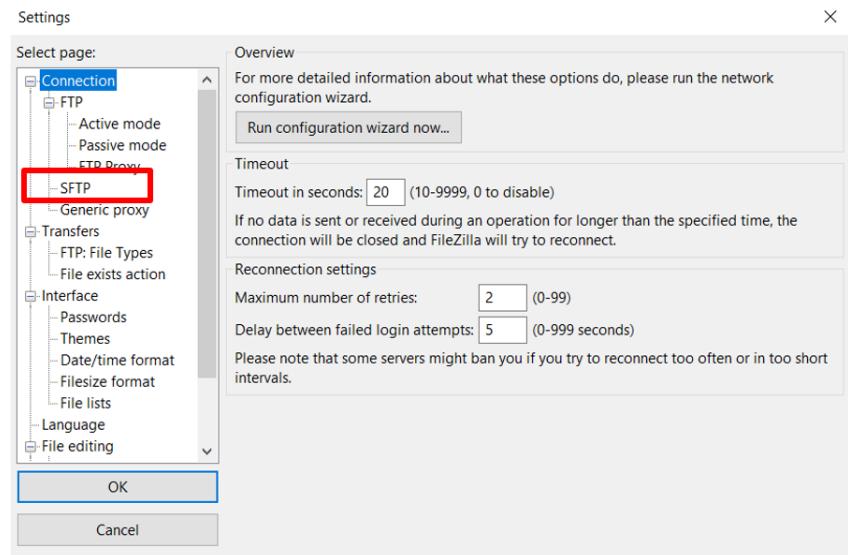
1. Open FileZilla Client.



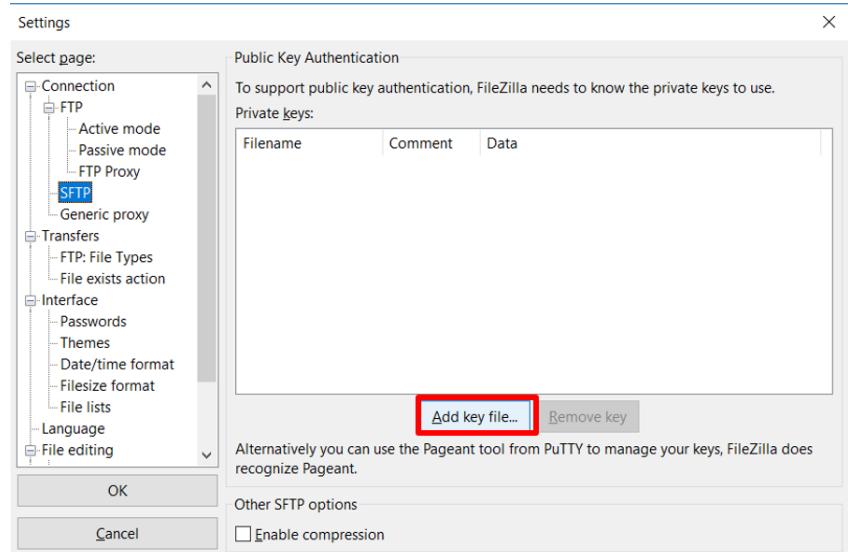
2. Click on Edit and then Settings.



3. Click on SFTP on the left hand side.

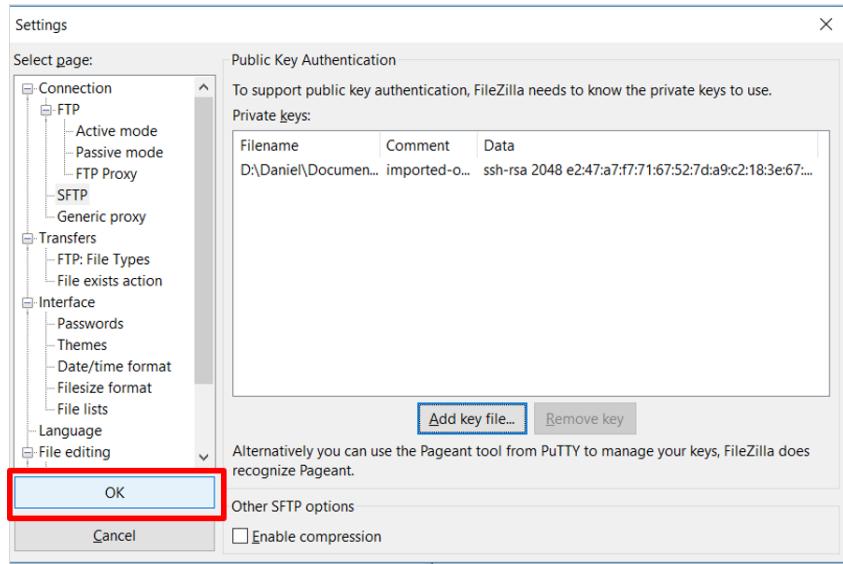


4. Click on 'Add key file...'



5. Select the key file that was generated using PuTTYgen.

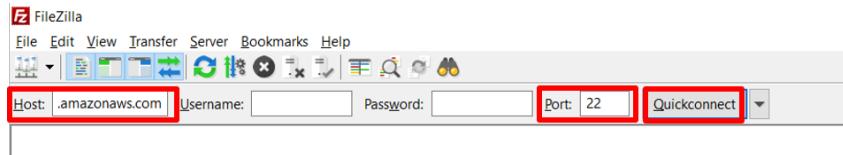
6. Click 'OK'.



7. Enter the following details:

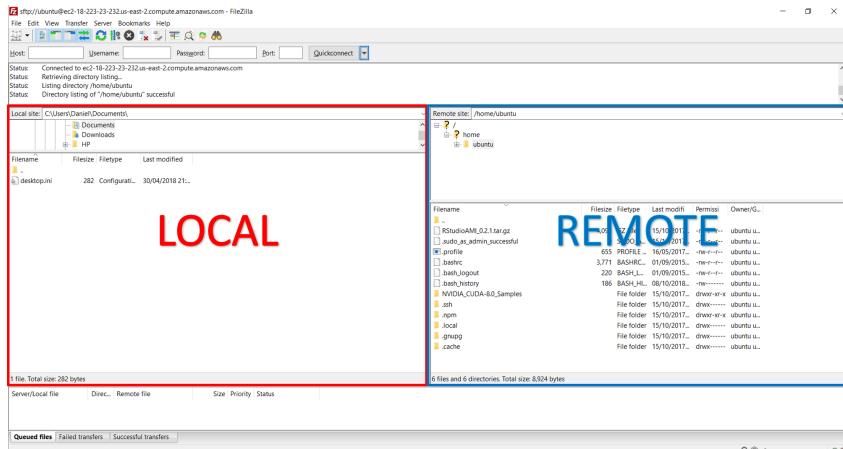
- (a) Host: 'ubuntu@###' where ### is the Public DNS (IPv4) from the EC2 Instance Dashboard.
- (b) Port: 22

8. Click 'Quickconnect'

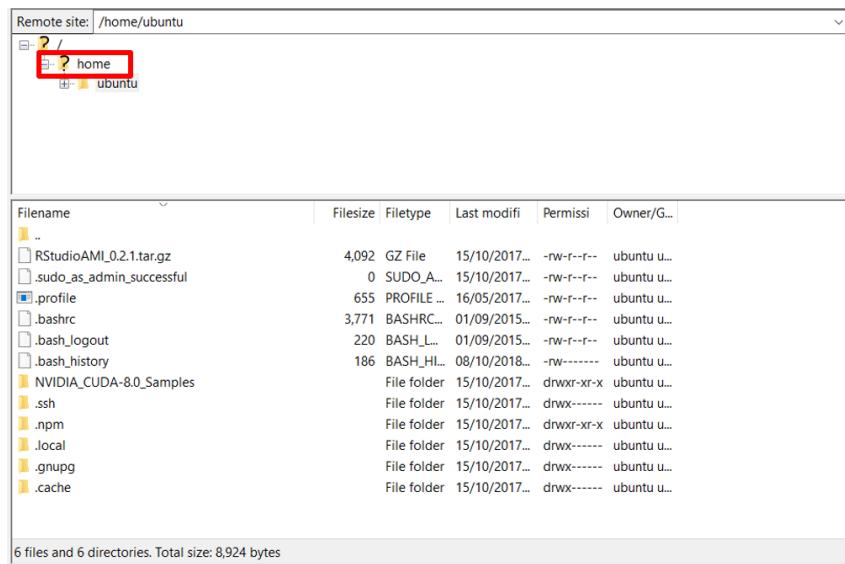


7.4 Adding Files for Use in R

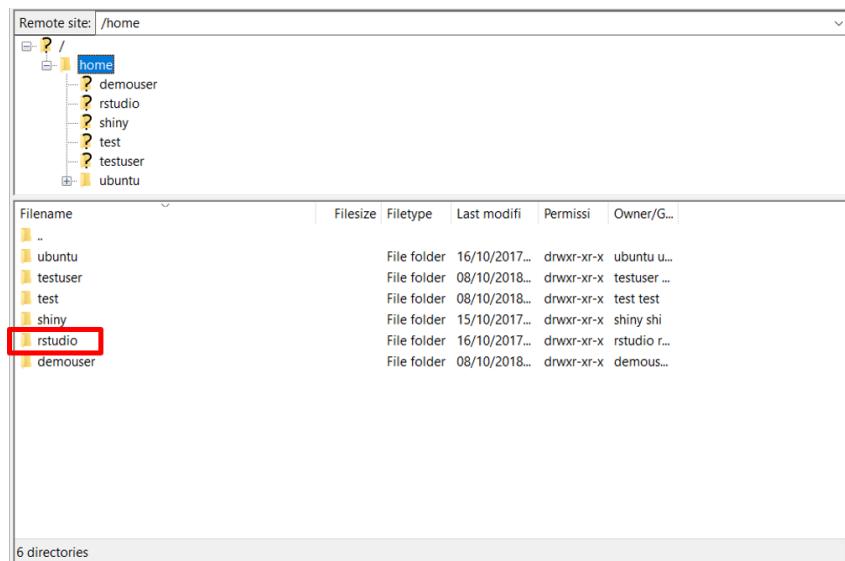
- 1. In FileZilla the left hand side displays your local directory and the right hand side displays the remote directory.



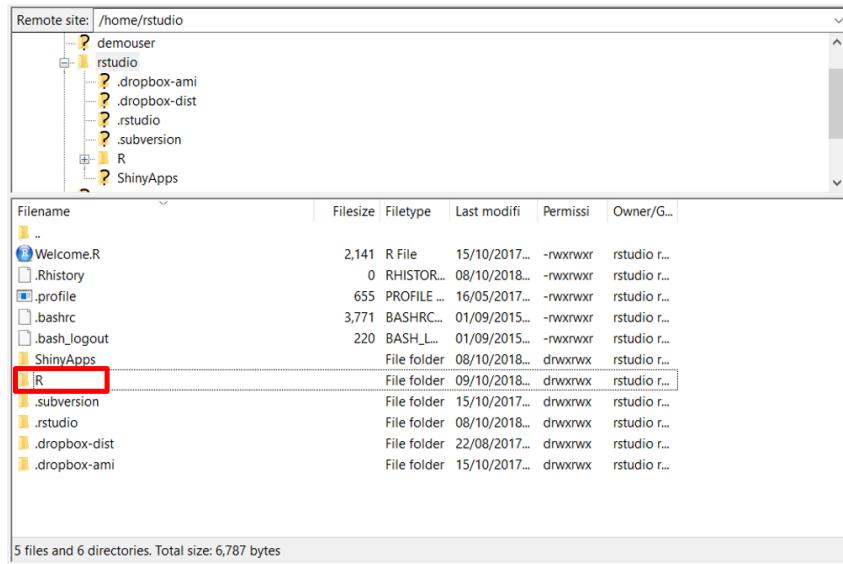
- 2. Click on 'home' on the Remote site.



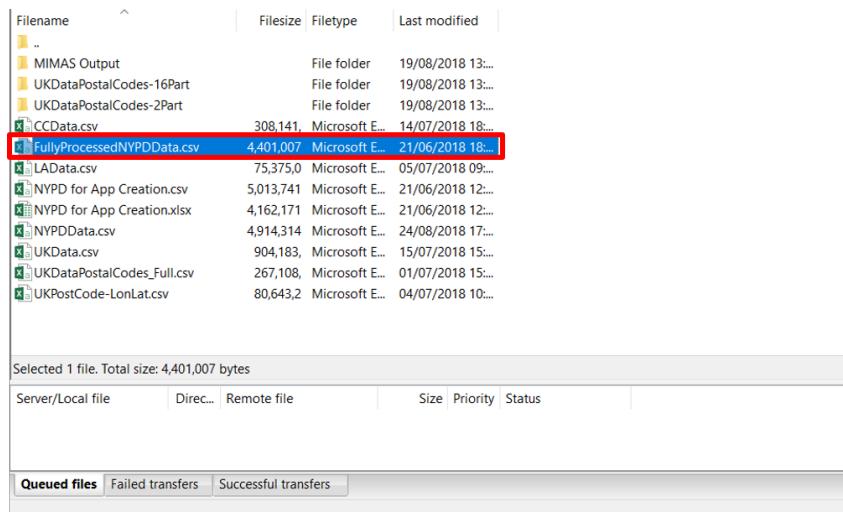
3. Then select the user you want to upload files for.



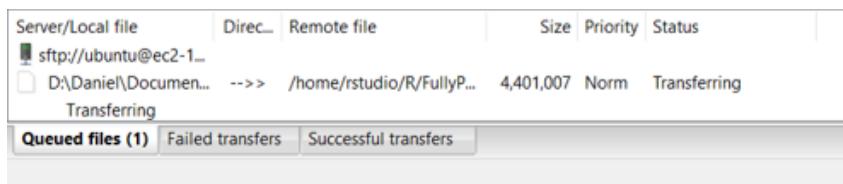
4. Choose the folder you wish to upload files to.



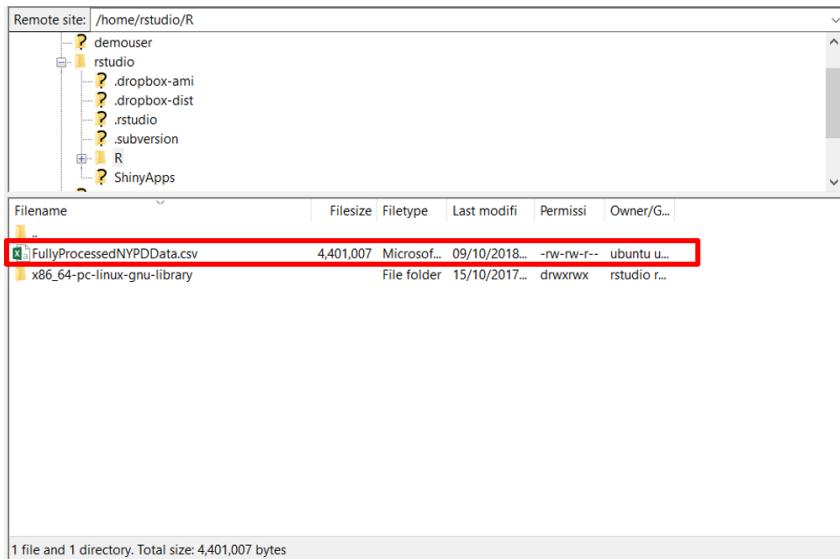
5. Then double click on the file you want to transfer.



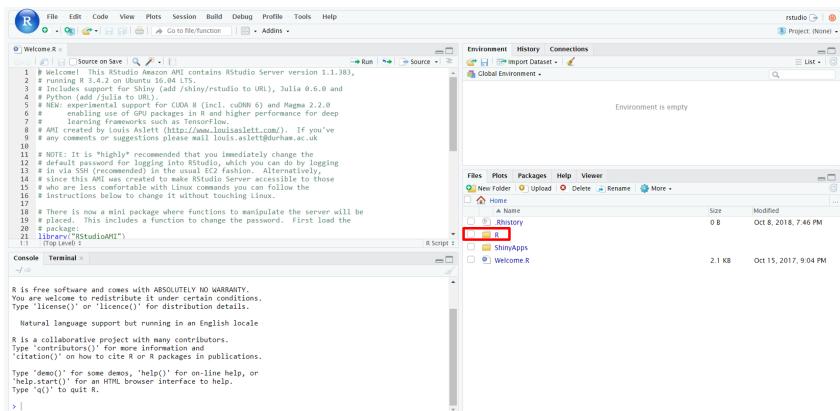
6. The transfer progress will be shown at the bottom of the FileZilla window.



7. Once the file has been successfully transferred close FileZilla.



8. Access the RStudio Server GUI as discussed above.



9. When the relevant user logs in they will see the uploaded files under the relevant folder. These files can then be used in R.

