

SETTING UP RSTUDIO NOTEBOOK FOR THE SCRNAMEQ WORKSHOP:

Navigate to the Sage Intranet->Science->Computing page and scroll to the following header. Using your Synapse credentials, click the indicated link (red arrow) to log into the AWS Service Catalog page.

The screenshot shows the Sage Intranet page with the Amazon Web Services logo. Below the header, there is a section titled "First, try the Service Catalog" which explains that the most common varieties of computing resources can be created through the AWS Service Catalog maintained by Sage. It mentions that users can log into sc.sageit.org with their Synapse account. Below this, a list of resource types is provided: EC2 instances, EC2 Ubuntu Linux, and S3 buckets. A red arrow points to the sc.sageit.org link, and a blue arrow points to the [EC2 instances](#) link.


Click “EC2: Ubuntu Linux with Notebook Software,” Launch Product, and follow the prompts. Detailed directions found at the link indicated by blue arrow, above, with the following specifics added:

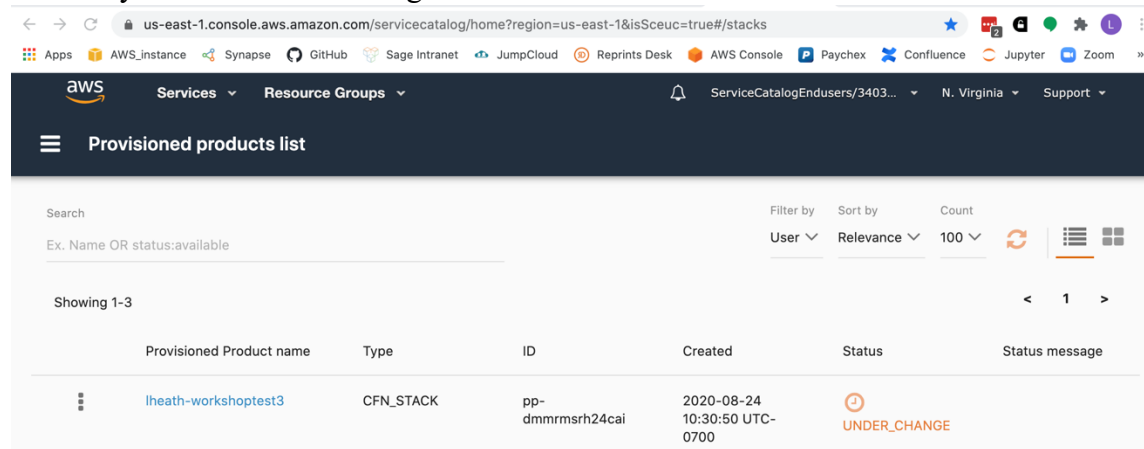
For EC2 Instance Type, scroll to “r4.xlarge.” Notebook type = Rstudio. Disk size = 16.

On TagOptions page, chose Project=amp-ad, Department=NDR.

The screenshot shows the AWS Service Catalog console. The "Products list" table is displayed with columns for Product name, Vendor, Owner, and Description. The product "EC2: Ubuntu Linux with Notebook Software" is highlighted with a blue circle. The table lists several products, including EC2: Linux, EC2: Ubuntu Linux with Notebook Software, EC2: Ubuntu Linux with Workflow Software, EC2: Windows, and S3: Private Encrypted Bucket.

Product name	Vendor	Owner	Description
EC2: Linux	Sage Bionetworks	Sage Bionetworks	This product builds one Linux EC2 instance, either Amazon Linux or Ubuntu
EC2: Ubuntu Linux with Notebook Software	Sage Bionetworks	Sage Bionetworks	This product builds one Linux EC2 instance using an Rstudio AMI based on Ubuntu and maintained by Sage Bionetworks.
EC2: Ubuntu Linux with Workflow Software	Sage Bionetworks	Sage Bionetworks	This product builds one Linux EC2 instance using an Ubuntu AMI with workflows software installed. Maintained by Sage Bionetworks.
EC2: Windows	Sage Bionetworks	Sage Bionetworks	This product builds one Microsoft Windows EC2 instance with Jumpcloud integration.
S3: Private Encrypted Bucket	Sage Bionetworks	Sage Bionetworks	This product builds an AWS S3 bucket encrypted with private access accessible from any source.

Go to the “provisioned products list” menu (found in the menu icon, upper left corner ) and wait for your notebook to change status.



us-east-1.console.aws.amazon.com/servicecatalog/home?region=us-east-1&isSceuc=true#/stacks

Services Resource Groups


ServiceCatalogEndUsers/3403... N. Virginia Support

Provisioned products list

Search
Ex. Name OR status:available

Filter by User Relevance Sort by Relevance Count 100

Showing 1-3

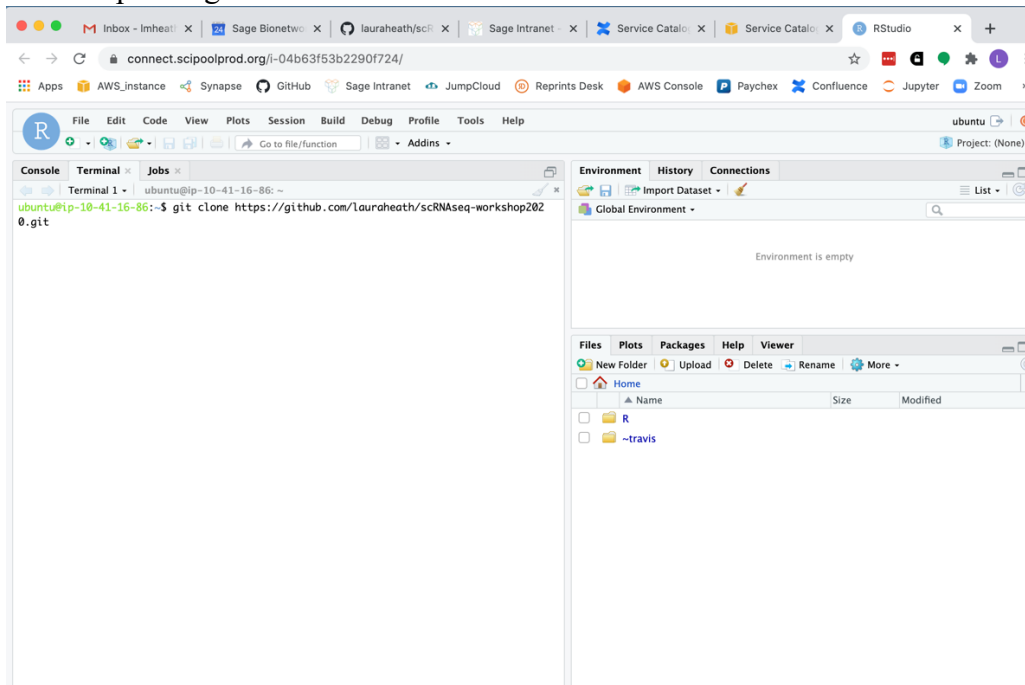
Provisioned Product name	Type	ID	Created	Status	Status message
lhealth-workshoptest3	CFN_STACK	pp-dmmrmsrh24cai	2020-08-24 10:30:50 UTC-0700	 UNDER_CHANGE	

When the Status says **Available**, click on the provisioned product name (in this example, [lhealth-workshoptest3](#)). Scroll down to the Outputs list and Click on the Value link next to NotebookConnectionURI. This will open a new Rstudio notebook in a new tab in your browser.

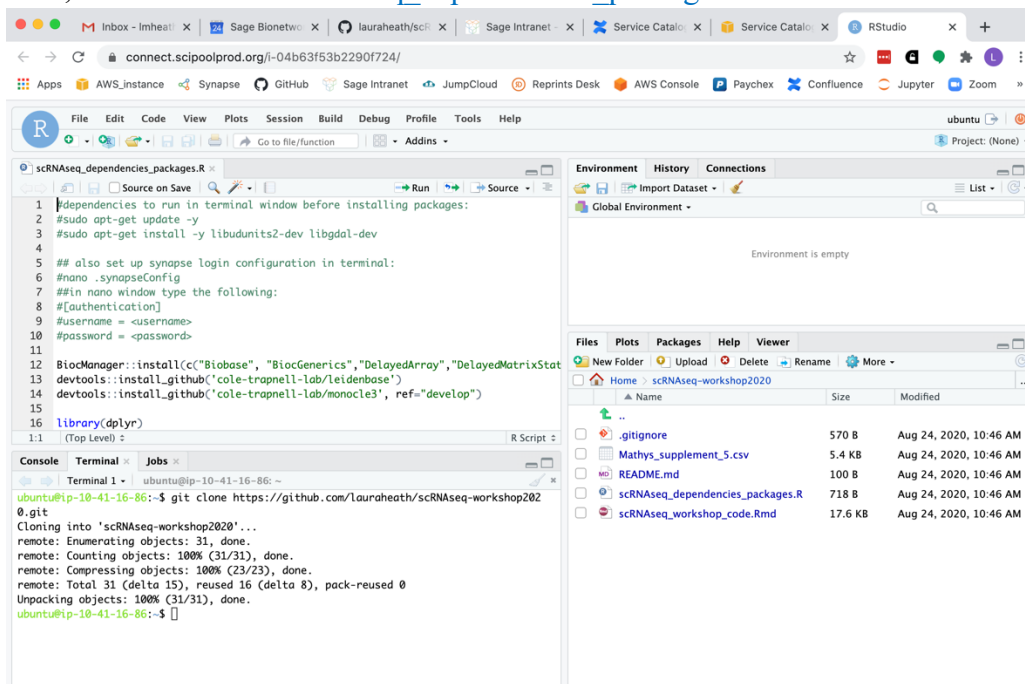
▼ Outputs:

Key	Value	Description
CloudformationStackARN	arn:aws:cloudformation:us-east-1:237179673806:stack/SC-237179673806-pp-dmmrmsrh24cai/8e600410-e62f-11ea-8a70-12695902b27f	The ARN of the launched Cloudformation Stack
EC2ConsoleURI	https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#Instances:search=i-04b63f53b2290f724	Check your instance status with this link to the ...
LinuxInstanceid	i-04b63f53b2290f724	The ID of the EC2 instance
LinuxInstancePrivateIpAddress	10.41.16.86	The IP Address of the EC2 instance
ConnectionInstructions	https://sagebionetworks.jira.com/wiki/spaces/IT/pages/996376579/Connect+to+Provisioned+Instances	Guidelines on connecting to instances
NotebookConnectionURI	https://connect.scipoolprod.org/i-04b63f53b2290f724/	Notebook server login page
EC2InstanceType	r4.2xlarge	The EC2 instance type
ConnectionURI	https://us-east-1.console.aws.amazon.com/systems-manager/session-manager/i-04b63f53b2290f724?region=us-east-1	Starts a shell session in the AWS Console

Go to the Rstudio Terminal (next to the Console tab, upper left) and clone the scRNA seq workshop repo with the following command: `git clone https://github.com/lauraheath/scRNAseq-workshop2020.git`



You will see the scRNAseq-workshop2020 folder added to your files (lower right panel). Click on it, then click on “[scRNAseq_dependencies_packages.R](#)”.



First, we need to install some dependencies. Copy line 2 (“`sudo apt-get update -y`” without the hashtag), paste into the terminal window prompt, and hit return; then do the same for line 3.

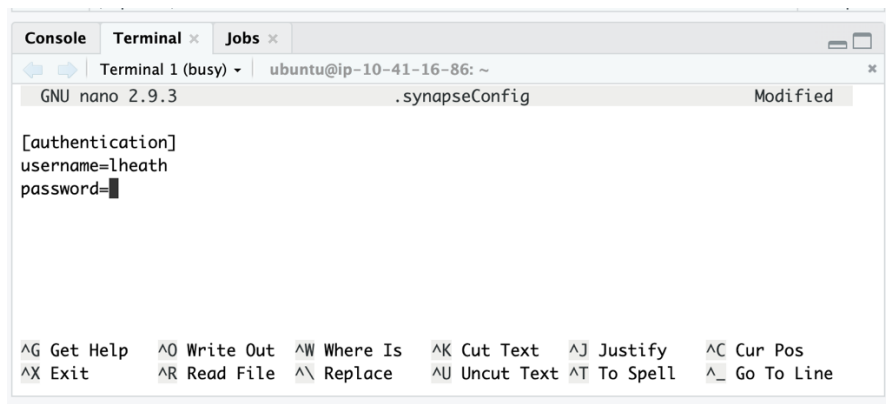
To enable easy Synapse login from your notebook, set up your login configuration as follows: after typing “nano .synapseConfig” in Terminal, a nano environment will appear. Type the following, line by line (no hashtags, don’t forget the brackets, and use your own Synapse login information:

[authentication]

username = yourSynapseUsername

password = yourSynapsePassword

Exit nano (control-X). Type ‘y’ when prompted to save the .synapseConfig file.



```
Terminal 1 (busy) | ubuntu@ip-10-41-16-86: ~
GNU nano 2.9.3 .synapseConfig Modified

[authentication]
username=lheath
password=

^G Get Help  ^O Write Out  ^W Where Is  ^K Cut Text  ^J Justify   ^C Cur Pos
^X Exit      ^R Read File  ^\ Replace   ^U Uncut Text ^T To Spell  ^_ Go To Line
```

Now install the necessary packages one line at a time (Lines 12-14 of scRNAseq_dependencies_packages.R) and install libraries (lines 16-20). Installing the packages will take ~35 minutes. Also, the monocle3 install will immediately prompt you to update existing packages (do update the existing packages by typing “1” when prompted).

Now you are ready to upload scRNA seq data from Synapse and use Monocle3 for trajectory analyses. Launch “scRNAseq_workshop_code.Rmd” and you are on your way!