

GETTING STARTED WITH RIVANNA AND JUPYTER NOTEBOOKS

Overview of Rivanna

Connecting to Rivanna with Open onDemand

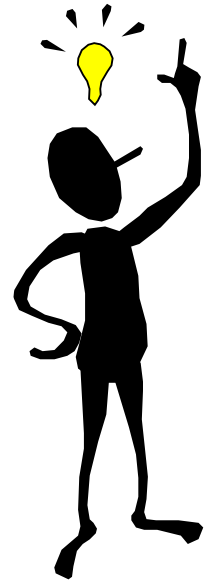
Copying Files

Starting a JupyterLab Session

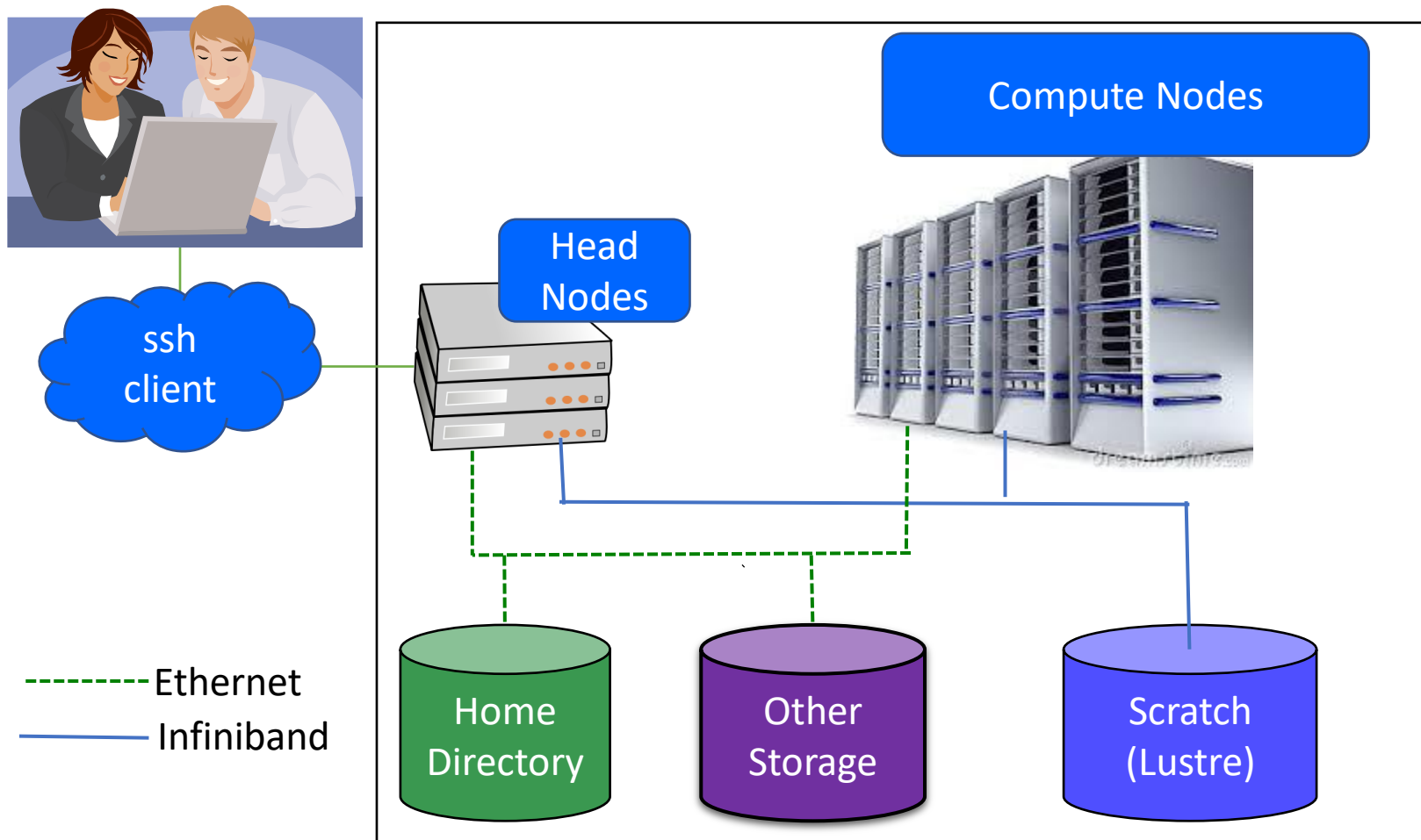
Running a Notebook

Terminology

- Rivanna is our high-performance cluster – basically. hundreds of computers networked together.
- Node
 - Basic building block of a cluster
 - Usually, a specialized computer
 - Two types of nodes:
 - **Head Node** – computer used for logging on and submitting jobs
 - **Compute Node** -- computer that does most of the work
- Core
 - an individual processor on a computer
 - The nodes on Rivanna have many cores (from 16 – 40)



Rivanna Overview



Connect to Rivanna

- You will be using our web portal, called Open onDemand, to connect to Rivanna.
- To access Open OnDemand, open your web browser and type <https://rivanna-portal.hpc.virginia.edu>

You will need to authenticate with Netbadge (i.e., “Netbadge” in).

- After you log in, you will see the Dashboard (shown on following slide).

Open onDemand Dashboard

UVA OpenOnDemand

Files


Jobs

Clusters

Interactive Apps

Help

Log Out

 UNIVERSITY
of VIRGINIA

Research Computing


OnDemand provides an integrated, single access point for all of your HPC resources.

Message of the Day

2019-10-08 Scratch Directory

RC system engineers will begin actively clearing /scratch files more than 90 days old **beginning 10/14/2019**. /scratch is intended as temporary storage (90 days maximum) for active work. It is not backed up and needs to be purged periodically in order to maintain a stable HPC environment. We encourage users to back up their important data. RC offers several low-cost storage options to researchers. For more information, visit <https://www.rc.virginia.edu/userinfo/rivanna/storage/>

powered by

OPEN  **nDemand**

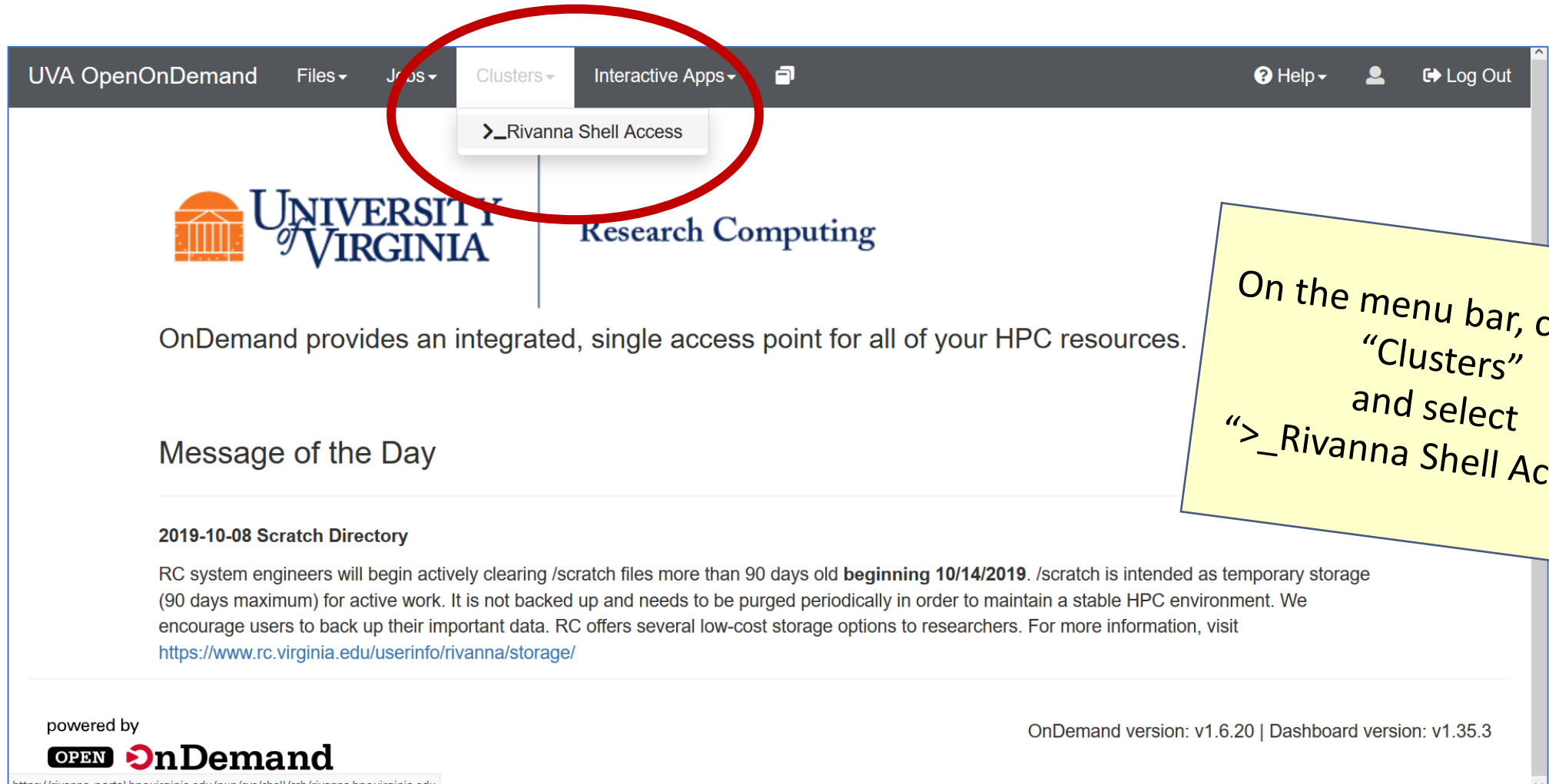
OnDemand version: v1.6.20 | Dashboard version: v1.35.3

*Congratulations!
When you see this page,
you are connected to
Rivanna*

Copying the Notebooks to your Account

- To get a copy of the sample Jupyter Notebooks in your account, you will open a shell (i.e., a terminal window) and use a command to copy the Notebooks (plus other necessary files) from a shared location to your scratch directory
- The next two slides will show how to do these steps.

Open a Shell



The screenshot shows the UVA OpenOnDemand dashboard. The top navigation bar includes 'UVA OpenOnDemand', 'Files', 'Jobs', 'Clusters', and 'Interactive Apps'. The 'Clusters' menu is highlighted with a red circle, and its sub-option '>_Rivanna Shell Access' is also highlighted. Below the navigation bar, the University of Virginia Research Computing logo is displayed. The main content area includes a message about HPC resources and a 'Message of the Day' section dated 2019-10-08 regarding the Scratch Directory. The footer shows the OpenOnDemand logo and version information.

UVA OpenOnDemand Files Jobs Clusters Interactive Apps Help Log Out

>_Rivanna Shell Access

UNIVERSITY of VIRGINIA Research Computing

OnDemand provides an integrated, single access point for all of your HPC resources.

Message of the Day

2019-10-08 Scratch Directory

RC system engineers will begin actively clearing /scratch files more than 90 days old **beginning 10/14/2019**. /scratch is intended as temporary storage (90 days maximum) for active work. It is not backed up and needs to be purged periodically in order to maintain a stable HPC environment. We encourage users to back up their important data. RC offers several low-cost storage options to researchers. For more information, visit <https://www.rc.virginia.edu/userinfo/rivanna/storage/>

powered by OPEN OnDemand

OnDemand version: v1.6.20 | Dashboard version: v1.35.3

On the menu bar, click on
"Clusters"
and select
">_Rivanna Shell Access"

The Shell

- A new tab will open with a black window and words like "Authorized Use Only!".

```
Last login: Sun Oct 25 22:44:16 2020 from 172.18.34.57
Authorized Use Only!
-bash-4.2$
```

- This is your shell. You can type Linux commands in the shell.

Command to Copy Files

- Type the following in the shell:

```
cp -r /project/rivanna-training/ML_with_Python /scratch/$USER
```

- This command will copy a folder and its contents to your scratch directory.

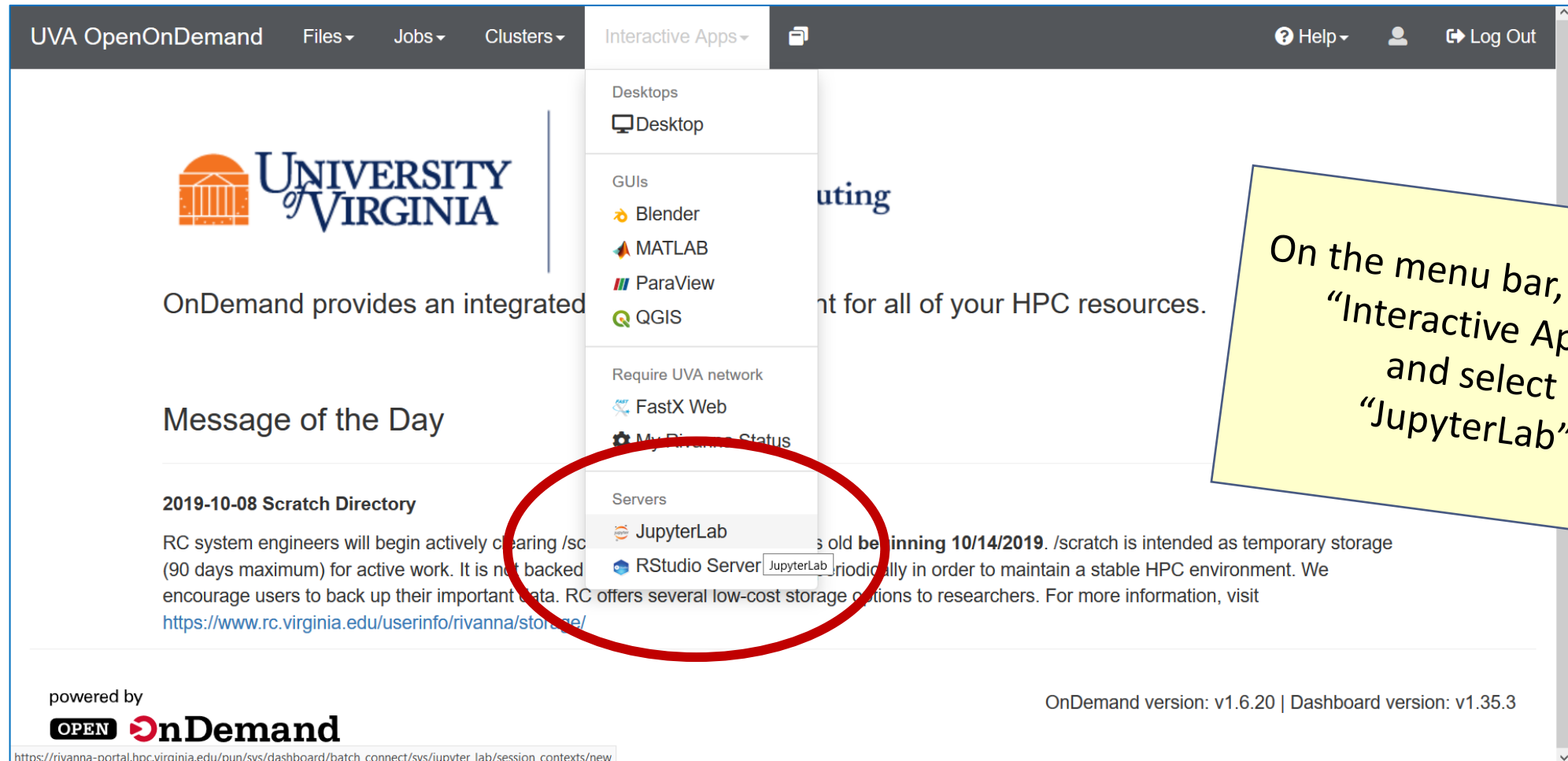
Exit out of the Shell

- You are now done with your shell. You can exit from it by closing the tab and clicking on “Leave page”.
- You should be back at the Open onDemand dashboard.
 - If not, click the tab on your browser that is labeled “Dashboard”.
 - If that fails, re-enter <https://rivanna-portal.hpc.virginia.edu> in you web browser.

Starting JupyterLab

- Before you can open up a Jupyter Notebook, you will need to request a JupyterLab session on Rivanna and provide some details for the resources that the Notebook needs. (For example, our sample Notebooks need to run on a GPU.)
- The next several slides will take you through the steps to start a JupyterLab session.

Select JupyterLab from Interactive Apps



The screenshot shows the UVA OpenOnDemand dashboard. The top navigation bar includes 'UVA OpenOnDemand', 'Files', 'Jobs', 'Clusters', 'Interactive Apps', 'Help', and 'Log Out'. The 'Interactive Apps' dropdown menu is open, showing categories: Desktops (Desktop), GUIs (Blender, MATLAB, ParaView, QGIS), Require UVA network (FastX Web, My Rivanna Status), and Servers (JupyterLab, RStudio Server). The 'JupyterLab' option is highlighted with a red circle. A yellow callout box on the right says: 'On the menu bar, click on "Interactive Apps" and select "JupyterLab"'. The main content area features the University of Virginia logo, a message about OnDemand integration, a 'Message of the Day', and a '2019-10-08 Scratch Directory' announcement. The footer includes 'powered by OPEN OnDemand' and version information: 'OnDemand version: v1.6.20 | Dashboard version: v1.35.3'.

UVA OpenOnDemand Files Jobs Clusters Interactive Apps Help Log Out

Desktops
Desktop

GUIs
Blender
MATLAB
ParaView
QGIS

Require UVA network
FastX Web
My Rivanna Status

Servers
JupyterLab
RStudio Server

On the menu bar, click on "Interactive Apps" and select "JupyterLab"

UNIVERSITY of VIRGINIA

OnDemand provides an integrated

Message of the Day

2019-10-08 Scratch Directory

RC system engineers will begin actively clearing /scratch (90 days maximum) for active work. It is not backed up. RC offers several low-cost storage options to researchers. For more information, visit <https://www.rc.virginia.edu/userinfo/rivanna/storage/>

powered by OPEN OnDemand

OnDemand version: v1.6.20 | Dashboard version: v1.35.3

https://rivanna-portal.hpc.virginia.edu/pun/sys/dashboard/batch_connect/sys/jupyter_lab/session_contexts/new

JupyterLab Session Selections (1/3)

Home / My Interactive Sessions / JupyterLab

Interactive Apps

Desktops

Desktop

GUIs

Blender

MATHEMATICA

MATLAB

ParaView

QGIS

Servers

Code Server

JupyterLab

This app will launch a Jupyter Lab server on one or more nodes.

Rivanna Partition

GPU

- Standard - (1-40 cores) Rivanna node in the standard partition.
- Instructional - Rivanna partition for Biocomplexity Institute and Institute.
- GPU - (1-28 cores) Rivanna node that has NVIDIA GPU.
- Dev - (1-8 cores) For short sessions (= 1 hour) with no SU charge; walltime is strictly limited to an hour.
- Instructional - (1-20 cores) Rivanna node in the instructional partition.
- [Learn More - Rivanna Queuing Policies](#)

Select the GPU Partition

To fill out the webform, follow the instructions on this and the two subsequent slides.

JupyterLab Session Selections (2/3)

Number of hours Set Number of hours to 2

2

Number of cores

1

Memory Request in GB (maximum 256G)

6

Work Directory Set Work Directory to SCRATCH

SCRATCH

Allocation (SUs)

rivanna-training

Optional: GPU type for GPU partition

default

Optional: Number of GPUs (1 ~ 4)

.

JupyterLab Session Selections (2/3)

Allocation (SUs)

rivanna-training

Set Allocation to rivanna-training

Optional: GPU type for GPU partition

default

Optional: Number of GPUs (1 4)

1

Set Number of GPUs to 1

Optional: Slurm Option

Optional: Group (for access to software or storage)

☐ I would like to receive an email when the session starts

Click on Launch

Launch

* The JupyterLab session data for this session can be accessed under the data root directory.

Note: During the workshop,
you will need to add
"--reservation=MLforPython"
in the Slurm Option Field

Wait for the Session to Start


JupyterLab (12492035) Queued

Created at: 2020-05-29 00:18:04 EDT

Time Requested: 3 hours

Session ID: 5763459f-60b3-4af3-a4f4-379d56a61354

Please be patient as your job currently sits in queue. The wait time depends on the number of cores as well as time requested.




JupyterLab (12492035) 1 node | 1 core | Running

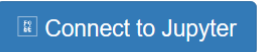
Host: >_udc-ba25-23

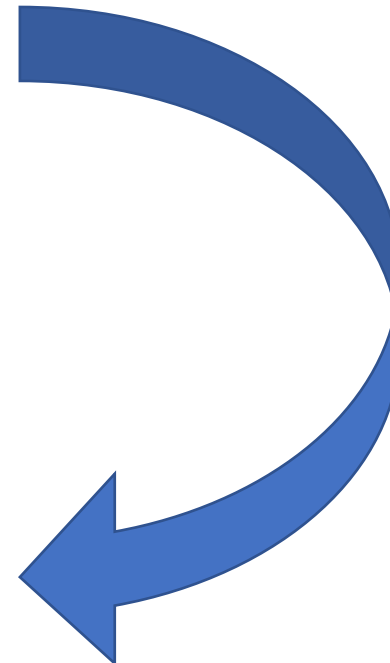
Created at: 2020-05-29 00:18:04 EDT

Time Remaining: 2 hours and 59 minutes

Session ID: 5763459f-60b3-4af3-a4f4-379d56a61354



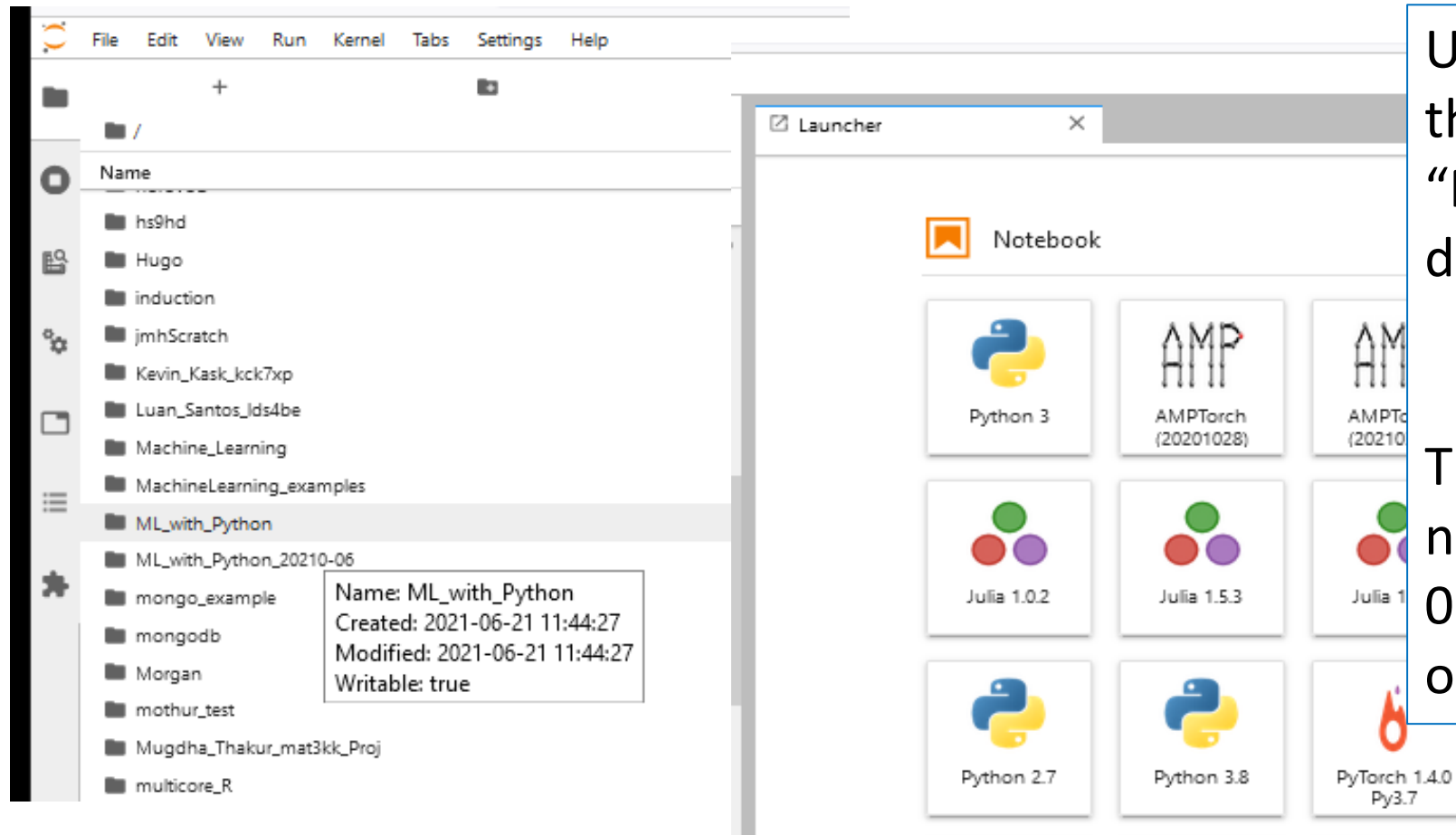




The screen will transition from a “Please be patient” statement to a “Connect to Jupyter” button.

Click on the “Connect to Jupyter” button.

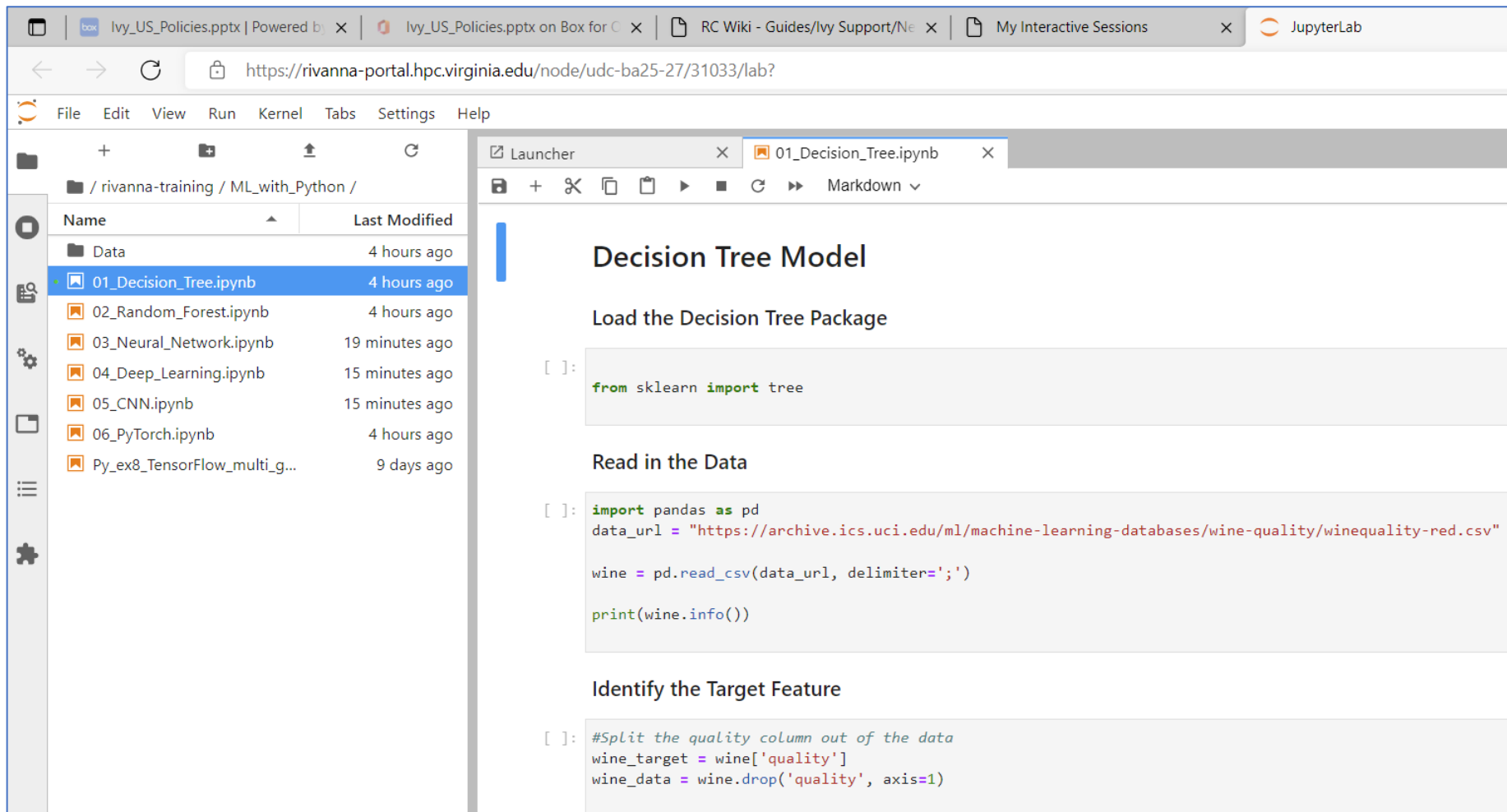
JupyterLab



Use the File Explorer to find the folder “ML_with_Python” and double-click on it.

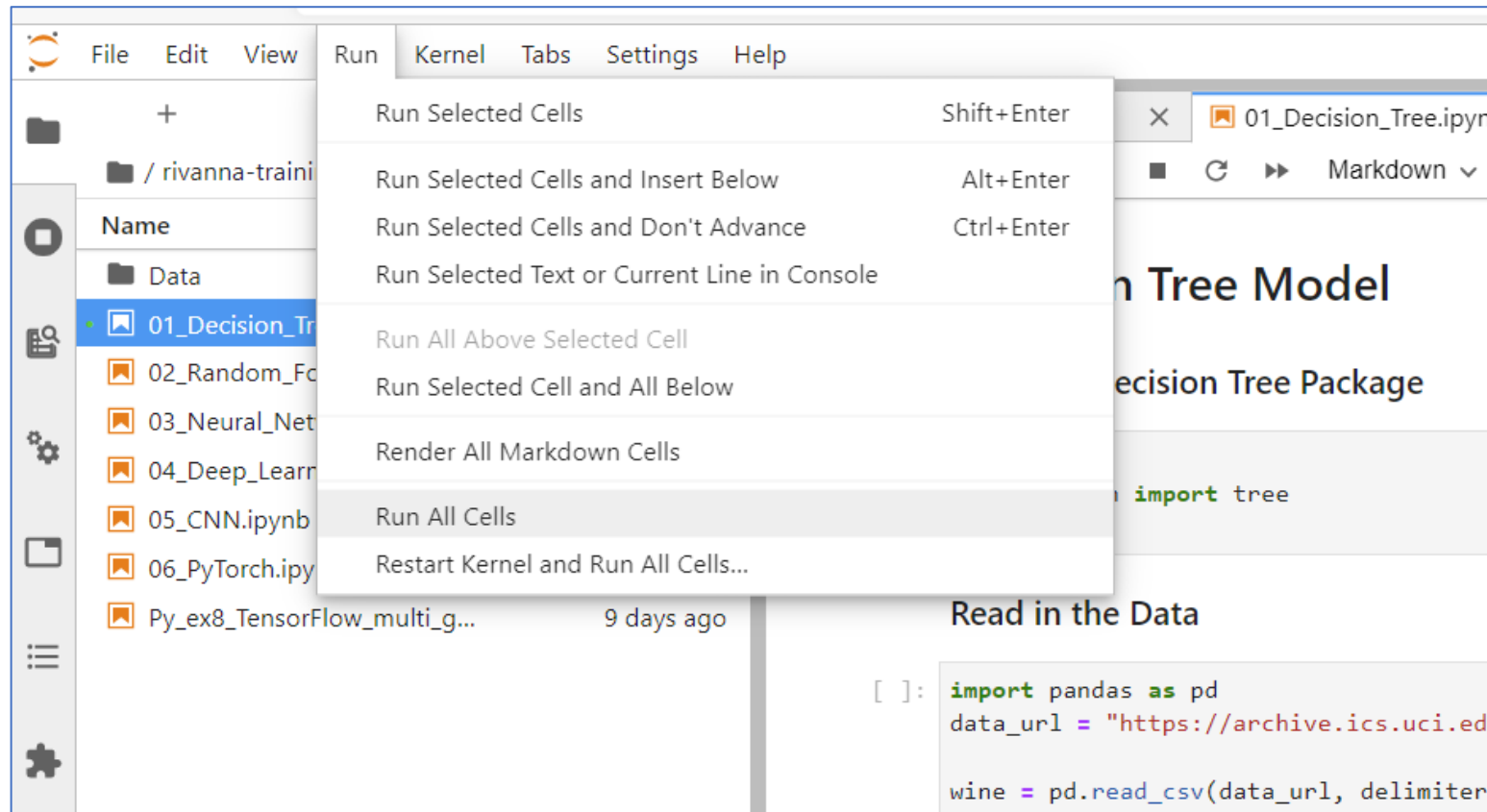
Then, double-click on the notebook file (e.g., 01_Decision_Tree.ipynb) to open it.

Viewing the Notebook



The screenshot displays a JupyterLab web interface in a browser. The address bar shows the URL <https://rivanna-portal.hpc.virginia.edu/node/udc-ba25-27/31033/lab?>. The interface includes a top menu bar with options like File, Edit, View, Run, Kernel, Tabs, Settings, and Help. On the left, a file browser shows the directory structure: /rivanna-training / ML_with_Python /, with a list of notebooks including '01_Decision_Tree.ipynb' (highlighted), '02_Random_Forest.ipynb', '03_Neural_Network.ipynb', '04_Deep_Learning.ipynb', '05_CNN.ipynb', '06_PyTorch.ipynb', and 'Py_ex8_TensorFlow_multi_g...'. The main workspace shows the '01_Decision_Tree.ipynb' notebook with a title 'Decision Tree Model'. The notebook content is divided into sections: 'Load the Decision Tree Package' with code `from sklearn import tree`; 'Read in the Data' with code `import pandas as pd`, `data_url = "https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-red.csv"`, `wine = pd.read_csv(data_url, delimiter=';')`, and `print(wine.info())`; and 'Identify the Target Feature' with code `#Split the quality column out of the data`, `wine_target = wine['quality']`, and `wine_data = wine.drop('quality', axis=1)`.

Running the Notebook



To run the Notebook, select *Run > Run All Cells*.

You will need to scroll through the Notebook to see all the results.

Ending the Session


JupyterLab (12492035) 1 node | 1 core | Running

Host: >_udc-ba25-23 Delete

Created at: 2020-05-29 00:18:04 EDT

Time Remaining: 2 hours and 59 minutes

Session ID: 5763459f-60b3-4af3-a4f4-379d56a61354

 [Connect to Jupyter](#)

When you are done with your Notebook, go back to the tab labeled “My Interactive Sessions” and click on the Delete button.

In the dialogue box that appears, click on the “Confirm” button.

Questions?

Questions about Rivanna or running
JupyterLab sessions can be sent to
hpc-support@virginia.edu

