Setting Up AIR-MS on Minerva OnDemand for Training Session I

September 19th 2025

Icahn School of Medicine at Mount Sinai

To prepare for AIRMS Training Session I, please follow the step-by-step setup guide. The guide will walk you through:

- Verifying prerequisites (accounts, software, connections)
- Accessing Minerva
- Launching our Training Jupyter Notebook

The Jupyter Notebook will serve as a follow-along resource during the training session, so it is important to confirm access in advance. We recommend that you complete this setup a few days before the session to ensure everything is working smoothly and to allow time for troubleshooting if needed.

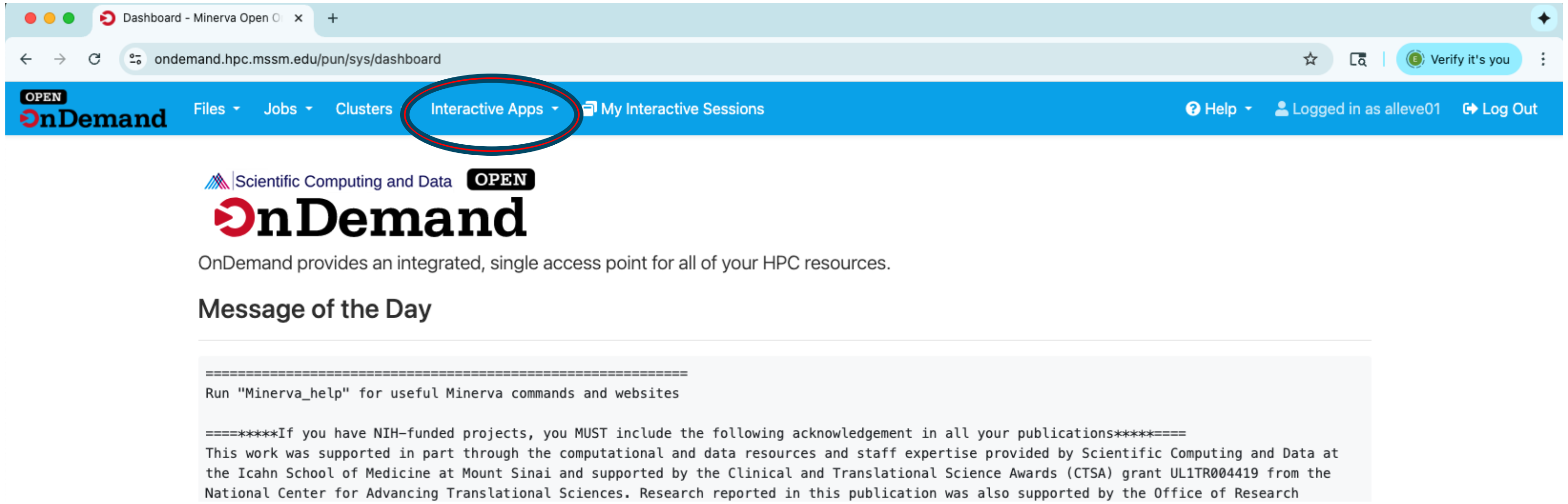
Pre-requisites

1. Possess a Mount Sinai School Network account
2. Request and obtain a Minerva account
3. Request and obtain access to AIR-MS MSDW DEID Dataset
4. Be either onsite on Mount Sinai Network (NOT Guest Wi-Fi) or connected to VPN

For more information on how to fulfill these steps visit our getting started page

Step 1: Launch a Jupyter Lab via OnDemand

1. Navigate to OnDemand via your browser
2. Sign in with your Mount Sinai credentials and password
3. Click on “Interactive Apps” and then select “Jupyter”

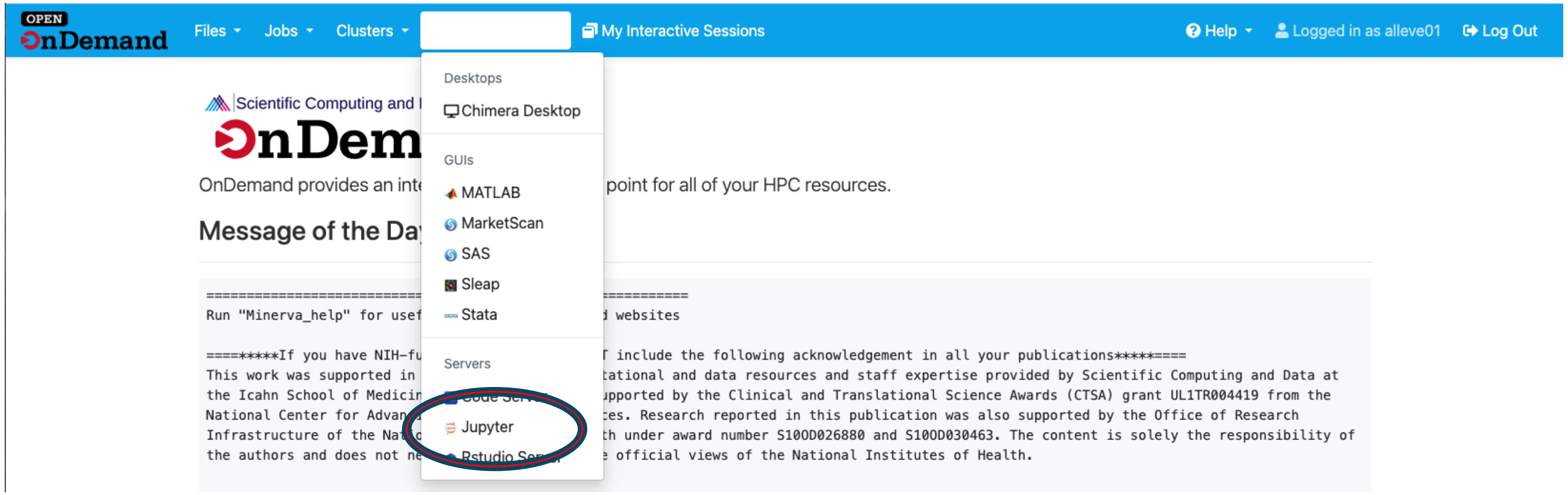


OnDemand provides an integrated, single access point for all of your HPC resources.

Message of the Day

```
================================================================
Run "Minerva_help" for useful Minerva commands and websites

====*****If you have NIH-funded projects, you MUST include the following acknowledgement in all your publications*****====
This work was supported in part through the computational and data resources and staff expertise provided by Scientific Computing and Data at
the Icahn School of Medicine at Mount Sinai and supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the
National Center for Advancing Translational Sciences. Research reported in this publication was also supported by the Office of Research
```



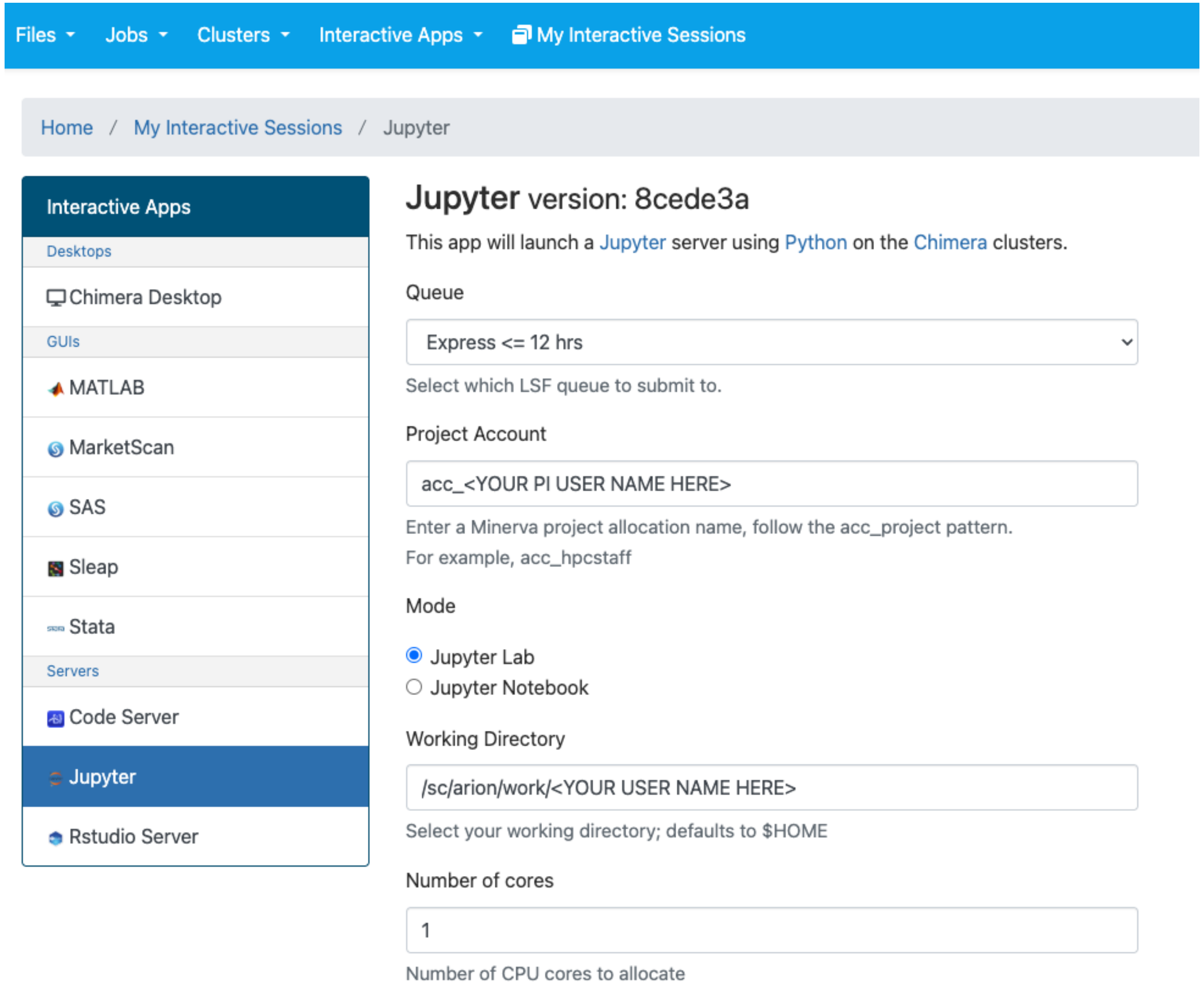
OnDemand provides an inte point for all of your HPC resources.

Message of the Day

```
Infrastructure of the Na                    th under award number S10OD026880 and S10OD030463. The content is solely the responsibility of
the authors and does not ne                  e official views of the National Institutes of Health.
```

Step 1: Launch a Jupyter Lab via OnDemand

1. Navigate to OnDemand via your browser
2. Sign in with your Mount Sinai credentials and password
3. Click on “Interactive Apps” and then select “Jupyter”
4. **Now fill out the form with the following and press “Launch”:**
   - **Queue:** Express <= 12 hrs
   - **Project Account:** The name of your Minerva project allocation, usually acc_<project name>. If you recently requested an account and do not have a project it might be something like acc_<your PI username>
   - **Mode:** Jupyter Lab
   - **Working Directory:** /sc/arion/work/<INSERT YOUR USER NAME HERE>
   - **Number of cores:** 1
   - **Memory request (in GB):** 1
   - **Number of hours:** 1 (you might want to select 2 during the training session and 1 for testing)
   - **Python version:** Python 3
   - **Extra Modules:** leave blank
   - **Reservation ID**: leave blank



Jupyter version: 8cede3a

This app will launch a Jupyter server using Python on the Chimera clusters.

Queue

Express <= 12 hrs

Select which LSF queue to submit to.

Project Account

acc_<YOUR PI USER NAME HERE>

Enter a Minerva project allocation name, follow the acc_project pattern.
For example, acc_hpcstaff

Mode

● Jupyter Lab
○ Jupyter Notebook

Working Directory

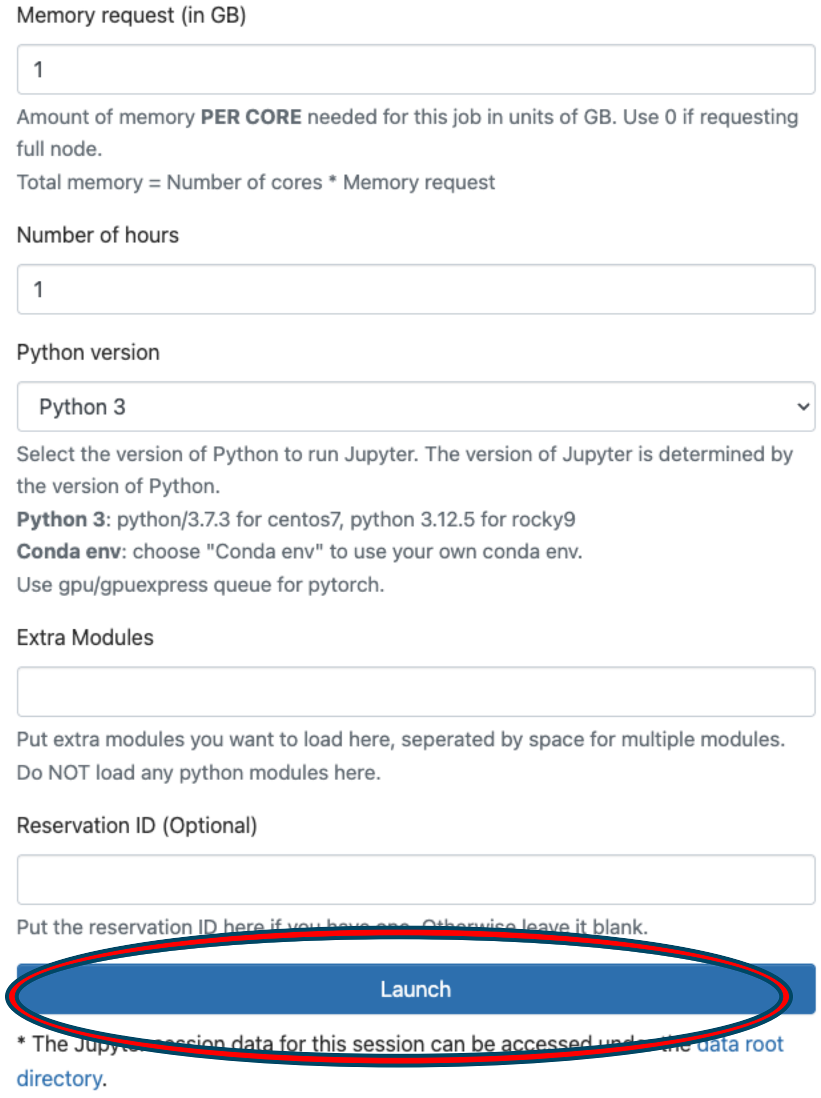
/sc/arion/work/<YOUR USER NAME HERE>

Select your working directory; defaults to $HOME

Number of cores

1

Number of CPU cores to allocate



Memory request (in GB)

1

Amount of memory **PER CORE** needed for this job in units of GB. Use 0 if requesting full node.
Total memory = Number of cores * Memory request

Number of hours

1

Python version

Python 3

Select the version of Python to run Jupyter. The version of Jupyter is determined by the version of Python.
Python 3: python/3.7.3 for centos7, python 3.12.5 for rocky9
Conda env: choose "Conda env" to use your own conda env.
Use gpu/gpuexpress queue for pytorch.

Extra Modules

Put extra modules you want to load here, seperated by space for multiple modules.
Do NOT load any python modules here.

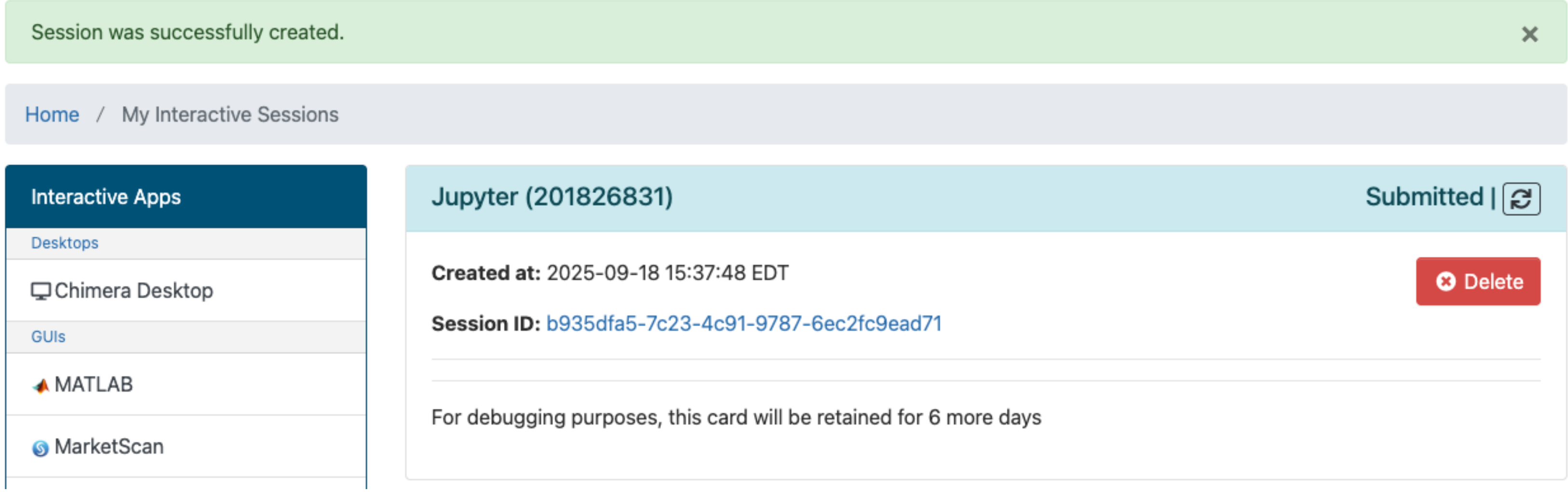
Reservation ID (Optional)

Put the reservation ID here if you have one. Otherwise leave it blank.

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

Step 1: Launch a Jupyter Lab via OnDemand

1. Navigate to OnDemand via your browser
2. Sign in with your Mount Sinai credentials and password
3. Click on “Interactive Apps” and then select “Jupyter”
4. Now fill out the form with the following and press “Launch”:
   - Queue: Express <= 12 hrs
   - Project Account: The name of your Minerva project allocation, usually acc_<project name>. If you recently requested an account and do not have a project it might be something like acc_<your PI username>
   - Mode: Jupyter Lab
   - Working Directory: /sc/arion/work/<INSERT YOUR USER NAME HERE>
   - Number of cores: 1
   - Memory request (in GB): 1
   - Number of hours: 1 (you might want to select 2 during the training session and 1 for testing)
   - Python version: Python 3
   - Extra Modules: leave blank
   - Reservation ID: leave blank
5. **Wait for the session to be in “Running” status and then press the “Connect to Jupyter” button**



Session was successfully created.

Home / My Interactive Sessions

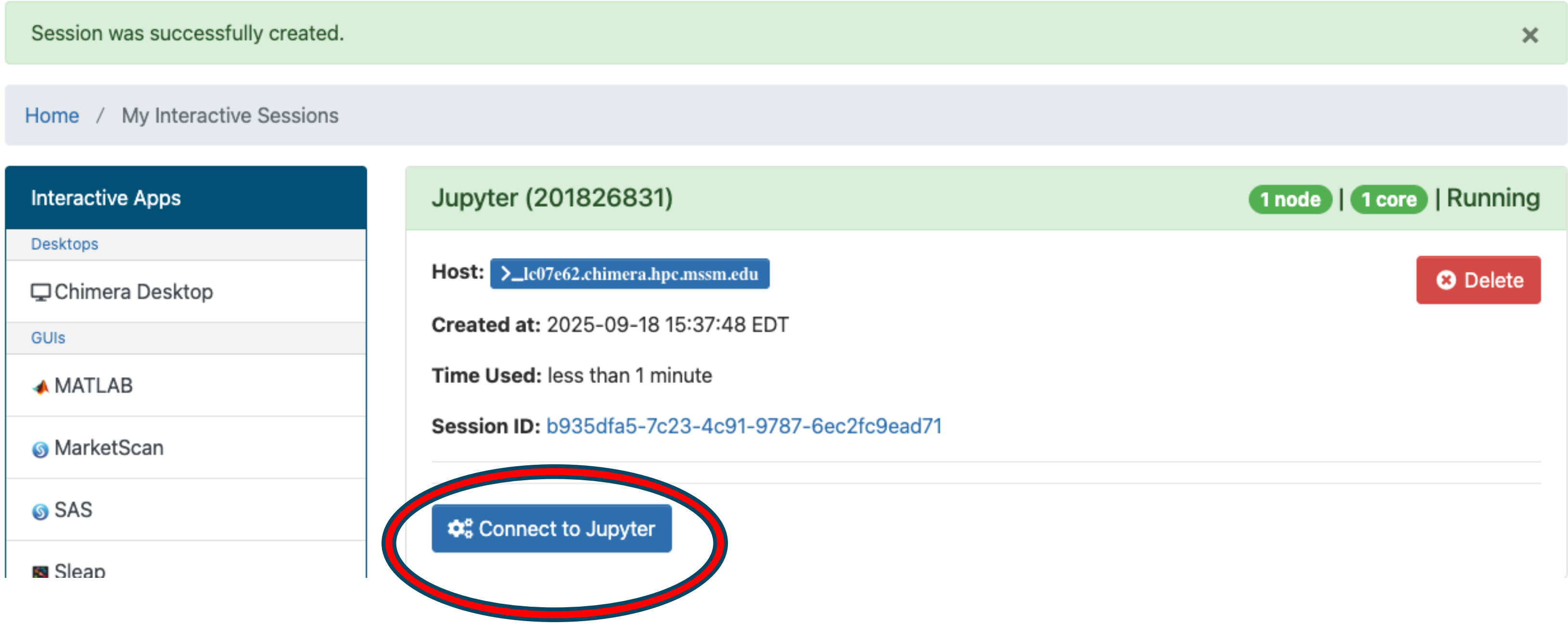
Jupyter (201826831)

Submitted |

Created at: 2025-09-18 15:37:48 EDT

Session ID: b935dfa5-7c23-4c91-9787-6ec2fc9ead71

For debugging purposes, this card will be retained for 6 more days



Session was successfully created.

Home / My Interactive Sessions

Jupyter (201826831)

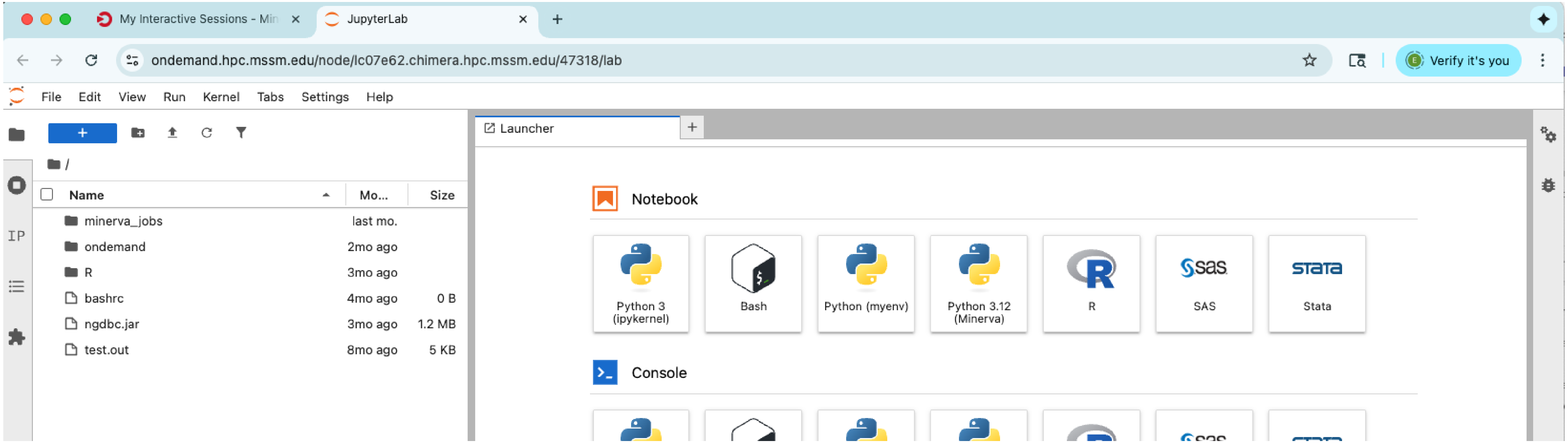
1 node | 1 core | Running

Host: >_lc07e62.chimera.hpc.mssm.edu

Created at: 2025-09-18 15:37:48 EDT

Time Used: less than 1 minute

Session ID: b935dfa5-7c23-4c91-9787-6ec2fc9ead71

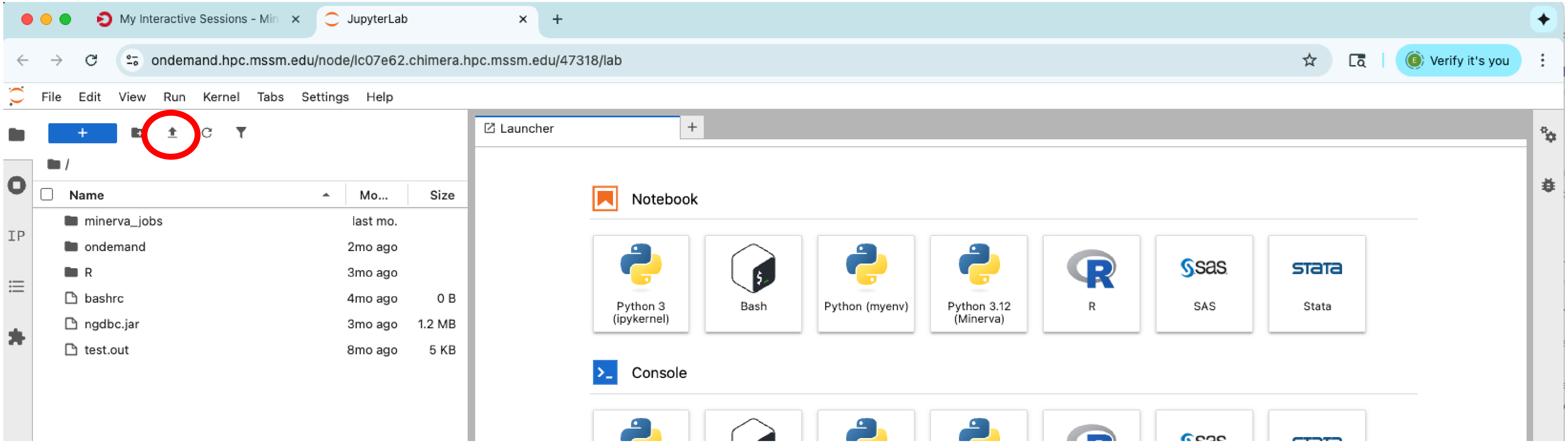


Name	Mo...	Size
minerva_jobs	last mo.	
ondemand	2mo ago	
R	3mo ago	
bashrc	4mo ago	0 B
ngdbc.jar	3mo ago	1.2 MB
test.out	8mo ago	5 KB

For help you can send an email to airms-support@mssm.edu

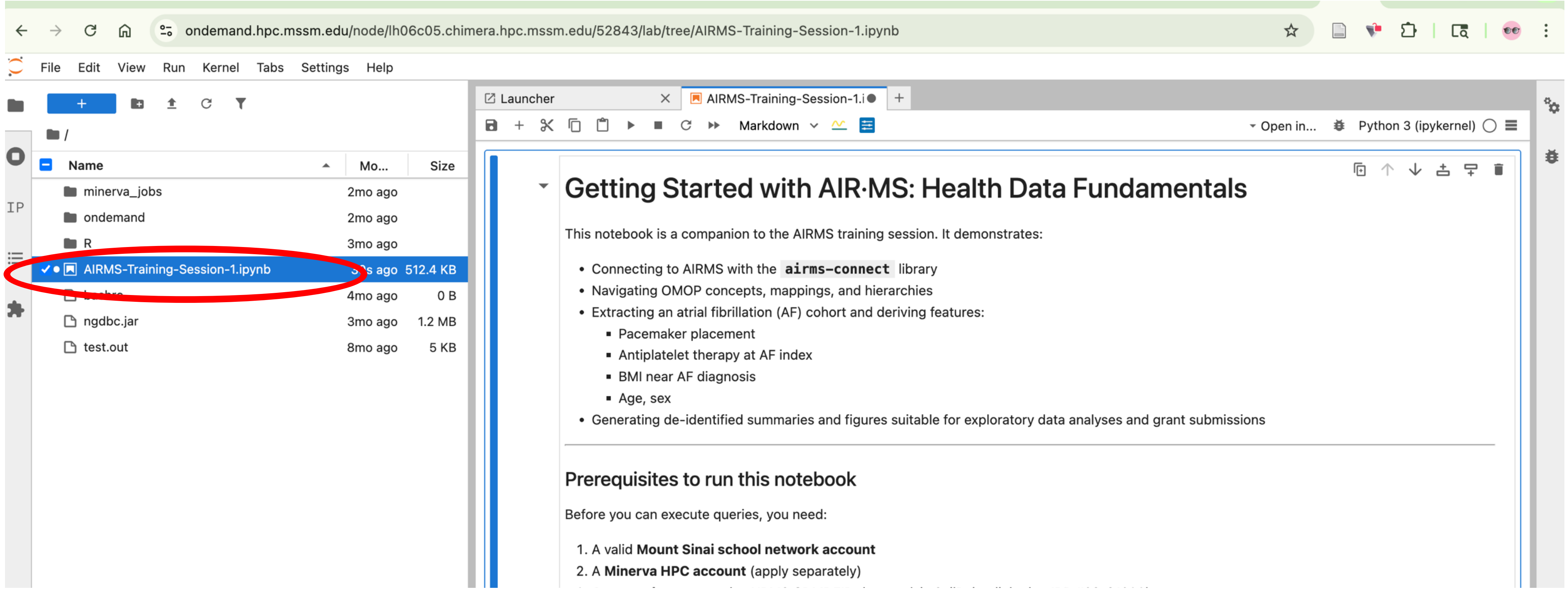
Step 2: Run the Jupyter Notebook

1. Download our example Jupyter Notebook at this link and drag & drop or upload (upper arrow icon button) it in your Jupyter Lab. Click on it once it appears on the left panel



Name	Mo...	Size
minerva_jobs	last mo.	
ondemand	2mo ago	
R	3mo ago	
bashrc	4mo ago	0 B
ngdbc.jar	3mo ago	1.2 MB
test.out	8mo ago	5 KB

For help you can send an email to airms-support@mssm.edu



Name	Mo...	Size
minerva_jobs	2mo ago	
ondemand	2mo ago	
R	3mo ago	
AIRMS-Training-Session-1.ipynb	s ago	512.4 KB
[illegible]	4mo ago	0 B
ngdbc.jar	3mo ago	1.2 MB
test.out	8mo ago	5 KB

Getting Started with AIR·MS: Health Data Fundamentals

This notebook is a companion to the AIRMS training session. It demonstrates:

- Connecting to AIRMS with the `airms-connect` library
- Navigating OMOP concepts, mappings, and hierarchies
- Extracting an atrial fibrillation (AF) cohort and deriving features:
  - Pacemaker placement
  - Antiplatelet therapy at AF index
  - BMI near AF diagnosis
  - Age, sex
- Generating de-identified summaries and figures suitable for exploratory data analyses and grant submissions

---

Prerequisites to run this notebook

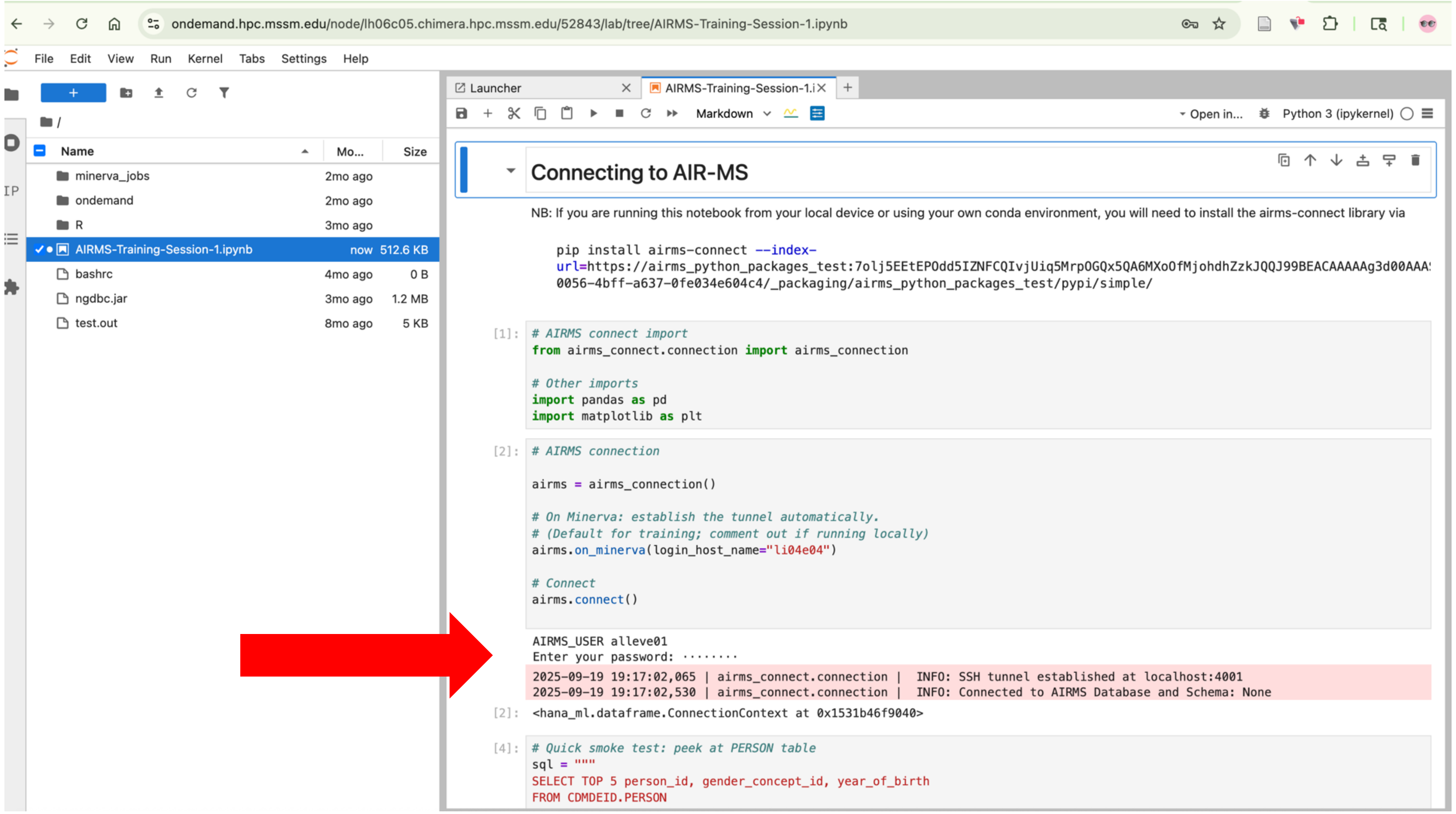
Before you can execute queries, you need:

1. A valid **Mount Sinai school network account**
2. A **Minerva HPC account** (apply separately)

For help you can send an email to airms-support@mssm.edu

Step 2: Run the Jupyter Notebook

1. Download our example Jupyter Notebook at this link and drag & drop or upload (upper arrow icon button) it in your Jupyter Lab. Click on it once it appears on the left panel
2. **Run the cells in the "Connecting to AIR-MS" section to connect to AIR-MS**
3. **You will be asked to input**
    - **AIRMS_USER – which corresponds to your Minerva username**
    - **Password - which is your Mount Sinai password (without VIP token)**
4. You are now connected to AIR-MS and should be able to run all the following cells



Connecting to AIR-MS

NB: If you are running this notebook from your local device or using your own conda environment, you will need to install the airms-connect library via

```
pip install airms-connect --index-
url=https://airms_python_packages_test:7olj5EEtEPOdd5IZNFCQIvjUiq5MrpOGQx5QA6MXoOfMjohdhZzkJQQJ99BEACAAAAAg3d00AAA
0056-4bff-a637-0fe034e604c4/_packaging/airms_python_packages_test/pypi/simple/
```

```
[1]: # AIRMS connect import
     from airms_connect.connection import airms_connection

     # Other imports
     import pandas as pd
     import matplotlib as plt
```

```
[2]: # AIRMS connection

     airms = airms_connection()

     # On Minerva: establish the tunnel automatically.
     # (Default for training; comment out if running locally)
     airms.on_minerva(login_host_name="li04e04")

     # Connect
     airms.connect()
```

```
AIRMS_USER alleve01
Enter your password: ········
2025-09-19 19:17:02,065 | airms_connect.connection |  INFO: SSH tunnel established at localhost:4001
2025-09-19 19:17:02,530 | airms_connect.connection |  INFO: Connected to AIRMS Database and Schema: None
[2]: <hana_ml.dataframe.ConnectionContext at 0x1531b46f9040>
```

```
[4]: # Quick smoke test: peek at PERSON table
     sql = """
     SELECT TOP 5 person_id, gender_concept_id, year_of_birth
     FROM CDMDEID.PERSON
```

Step 2: Run the Jupyter Notebook

1. Download our example Jupyter Notebook at this link and drag & drop or upload (upper arrow icon button) it in your Jupyter Lab. Click on it once it appears on the left panel
2. Run the cells in the “Connecting to AIR-MS” section to connect to AIR-MS
3. You will be asked to input
   - AIRMS_USER – which corresponds to your Minerva username
   - Password - which is your Mount Sinai password (without VIP token)
4. **You are now connected to AIR-MS and should be able to run all the following cells**

Thank you.

For help, please reach out to airms-support@mssm.edu

For additional resources visit:

- Our Researcher Tutorials (Accessible on Mount Sinai Network or VPN)
- Our Website

Icahn School
of Medicine at
Mount
Sinai