Jupyter Notebooks Running Log

As difficult as you would expect

Rob Donald

Wednesday 17 January 2024

Table of contents

1 Executive Summary

A running log to support the understanding of how to view and use Jupyter Notebooks (python).

Note this is all done on an M2 MacBook Pro, so how it works on older Intel based Macs and Windows is unknown.

2 Background

In the mixed world of data science you often run across a thing called a "Jupyter Notebook" which is a python version of a `.Rmd` or `.qmd` file, i.e. it mixes text with code.

I think it may have been called an "IPython Notebook" in the past ¯_(ツ)_/¯.

This R project was motivated by a need to try to at least *read* a `.ipynb` file even if you can't get it to run.

This has been asked about recently (2023-10-07) on the Posit Community site at:

- https://community.rstudio.com/t/is-there-a-plugin-to-view-jupyter-notebooks-ipynb/174723

But with no answer.

As you will see it all seems a bit tedious but I'll admit to a bias on all this because I still find setting up a python environment *so* much more awkward that an RStudio project.

3 2024-01-17

Setting up the project.

3.1 Install Python

This is now pretty straightforward. For every python project you are are going to investigate you need to set up a virtual environment so that you can isolate things.

1. Create a new R Project
2. Set up a running log file
3. From the terminal

   - a) virtualenv Some_Sensible_Name_py_env
   - b) source Some_Sensible_Name_py_env/bin/activate
   - c) pip install numpy pandas matplotlib

This will get you a minimal environment. You can have a look at the file `Venv_Setup_20240117_01.txt` in this repo to see the sort of thing that will go screaming across the Rstudio terminal output as you run the above steps.

4. Quit and Restart the R Project. You should now see on the Rstudio terminal tab that the virtual environment is running. For this project it looks like this:

   - `(JupyterNB_py_env) Taylor:JupyterNotebooks rob$`

5. Install any other python packages you need. I usually install `scipy` at least

   - pip install scipy

You can have a look at the top of file `Venv_Setup_20240117_02.txt` in this repo to see the expected output.

3.2 Install Jupyter Notebook

Now we get into the specifics of this project. The Jupyter project at:

- https://jupyter.org/

gives you some background. It is a major competitor to RStudio and their Quarto and R Shiny systems as you can see from all the overlap.

For my purposes all I want to do is read a `.ipynb` file so I just need the 'Jupyter Notebook' application which is the classic interface:

- https://jupyter.org/install

Still at the Rstudio terminal prompt you just install it as a python package.

- pip install notebook

This will install *lots and lots* of things, you can see the expected output in the file `Venv_Setup_20240117_02.txt` in this repo.

This is an application which runs up a small local web server to read the `.ipynb` files.

3.2.1 Starting Jupyter Notebook

Remember this is an application so you start it from the RStudio terminal tab with:

- $ jupyter notebook

After it settles down you will see some log messages like this

```
[I 2024-01- ... ServerApp] Jupyter Server 2.12.5 is running at:
[I 2024-01- ... ServerApp] http://localhost:8888/tree?token=ca163 ...
[I 2024-01- ... ServerApp] http://127.0.0.1:8888/tree?token=ca163 ...
[I 2024-01- ... ServerApp] Use Control-C to stop this server and
shut down all kernels (twice to skip confirmation).
```

It starts up a web page in your default browser which you can now try to use to read the .ipynb file you are interested in. See Section 3.3

3.2.2 Stopping Jupyter Notebook

As it says above, use Control-C **in the RStudio terminal tab** to stop the server. If you do it once it comes up asking for confirmation. If you are not quick with the 'y' at the prompt it will restart ¯_(ツ)_/¯.

```
^C
[I 2024-01- ... ServerApp] interrupted
[I 2024-01- ... ServerApp] Serving notebooks from local directory:  /Users/...
    1 active kernel
    Jupyter Server 2.12.5 is running at:
    http://localhost:8888/tree?token=ca163 ...
    http://127.0.0.1:8888/tree?token=ca163 ...
Shutdown this Jupyter server (y/[n])?
[I 2024-01- ... ServerApp] No answer for 5s:
[I 2024-01- ... ServerApp] resuming operation...
```

Assuming you answer 'y' it will shutdown with a message like this:

```
Shutdown this Jupyter server (y/[n])? y
[C 2024-01- ... ServerApp] Shutdown confirmed
[I 2024-01- ... ServerApp] Shutting down 5 extensions
(JupyterNB_py_env) Taylor:JupyterNotebooks rob$
```

Note that just closing the web page it started in Section 3.2.1 will ***not*** stop the local web server.

3.3 Using Jupyter Notebook

The thing I have found is that the files you want to look at need to be under the project root for the Jupyter Notebook server. I can run with that.

There is a big hint with that if you try to use the file 'open' on the web page.
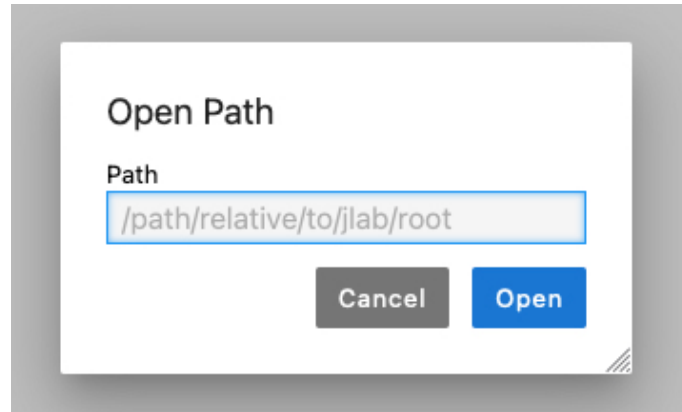

Figure 1: Reading a file

It says *relative to jlab root* which I take to mean the web server root which starts at your R Project's top level folder.

References