

Bonus

```
root@p930026096:~# docker run -d -p 8888:8888 -p 4040:4040 -p 4041:4041 -p 4042:4042 jupyter/pyspark-notebook 762837b5b1bcac6e75a4699ca15ed26e83cfe13a64151df6d5cf2d336dda6f0e
```

Run a docker container by specific ports

```
root@p930026096:~# docker ps -a
CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS
762837b5b1bc   jupyter/pyspark-notebook           "tini -g -- start-no-   6 seconds ago Up 5 seconds   0.0.0.0:4040-4042->4040-4042/tcp, :::4040-4042->4040-4042/tcp, 0.0.0.
0:8888->8888/tcp, :::8888->8888/tcp
81e3cd4668db   jupyter/pyspark-notebook           "tini -g -- -p 8888;-   16 hours ago   Exited (127) 16 hours ago
jupyter
43f153e36467   jupyter/pyspark-notebook           "tini -g -- start-no-   19 hours ago   Exited (255) 16 hours ago
0:49153->8888/tcp, :::49153->8888/tcp
ca77fec5f91    jupyter/pyspark-notebook           "tini -g -- start-no-   25 hours ago   Exited (0) 25 hours ago
gifted_ramanujan
```

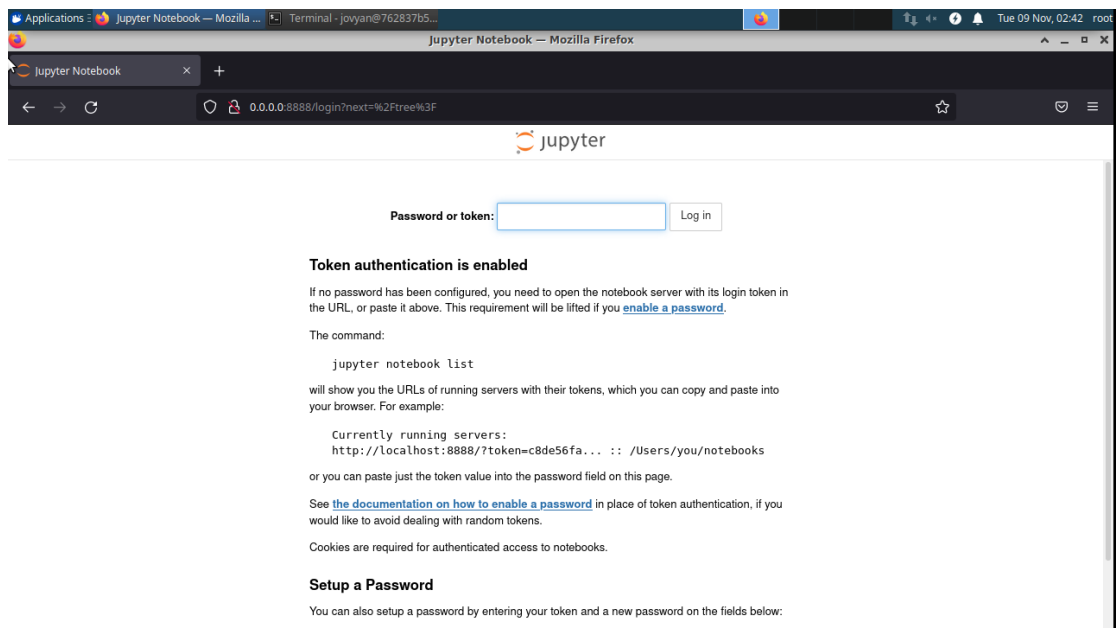
Check the container status, the first container is what we ran.

```
root@p930026096:~# docker exec -it 762837b5b1bc bash
(base) jovyan@762837b5b1bc:~$ jupyter notebook list
```

Copy the container ID and execute the container command line.

```
(base) jovyan@762837b5b1bc:~$ jupyter notebook list
Currently running servers:
http://0.0.0.0:8888/?token=ab05db0147edbad6ad0bef9842dc2229a8a2f44d68ba8e75 :: /home/jovyan
(base) jovyan@762837b5b1bc:~$
```

Type the command “jupyter notebook list” and it will show a series of token of website “http://0.0.0.0:8888/”



Applications : Jupyter Notebook — Mozilla Firefox | Terminal - jovyan@762837b5...

Jupyter Notebook — Mozilla Firefox

Jupyter Notebook

← → ↻ 0.0.0.0:8888/login?next=%2Ftree%3F

jupyter

Password or token: Log in

Token authentication is enabled

If no password has been configured, you need to open the notebook server with its login token in the URL, or paste it above. This requirement will be lifted if you [enable a password](#).

The command:

```
jupyter notebook list
```

will show you the URLs of running servers with their tokens, which you can copy and paste into your browser. For example:

```
Currently running servers:
http://localhost:8888/?token=c8de56fa... :: /Users/you/notebooks
```

or you can paste just the token value into the password field on this page.

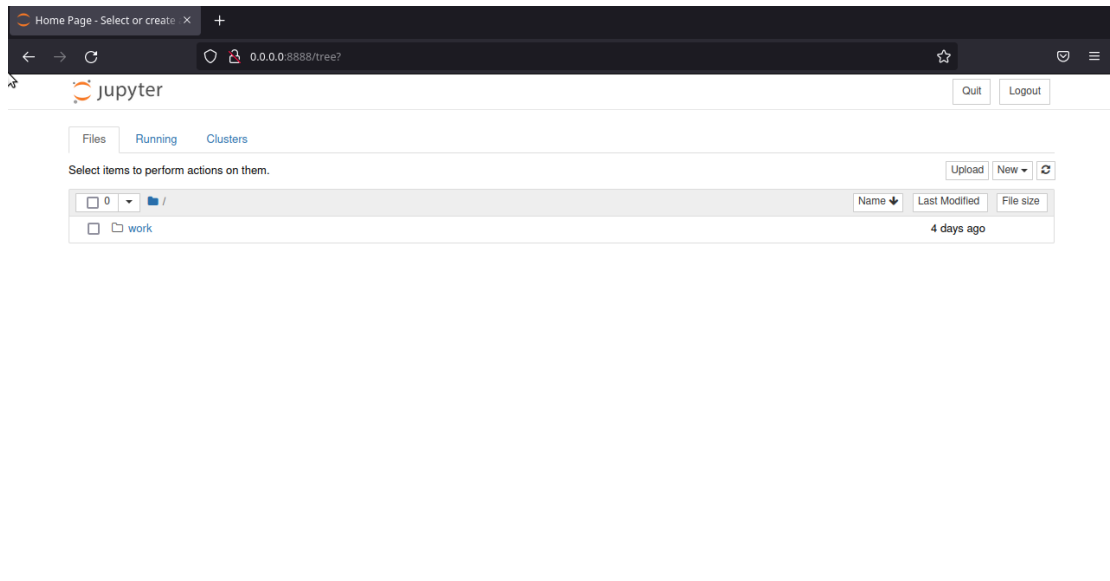
See [the documentation on how to enable a password](#) in place of token authentication, if you would like to avoid dealing with random tokens.

Cookies are required for authenticated access to notebooks.

Setup a Password

You can also setup a password by entering your token and a new password on the fields below:

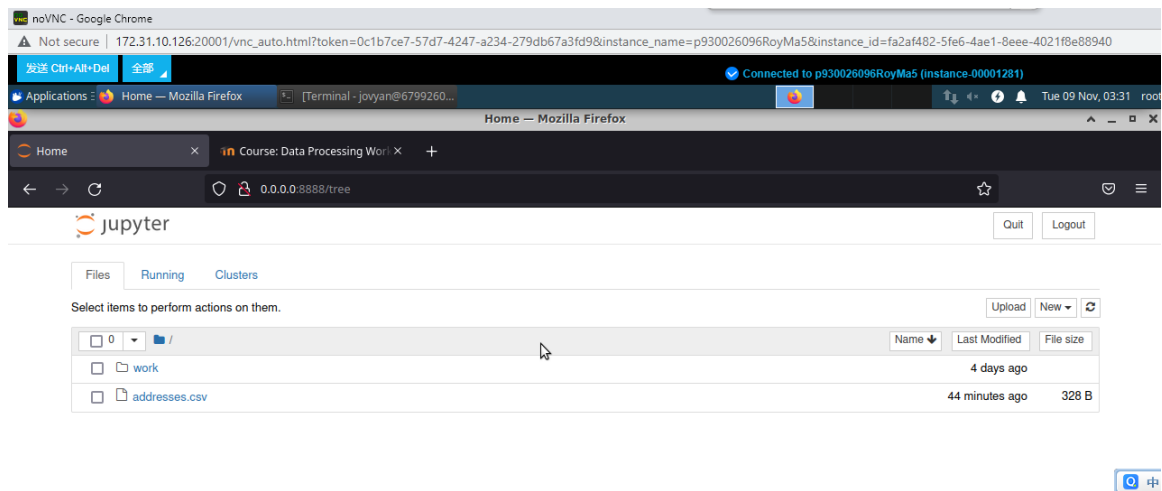
Open the website by web browser and copy the token to log in the jupyter notebook.



We found that currently we can create a python file but we didn't have the .csv file.

```
root@p930026096:~# docker cp ~/addresses.csv 67992604ce12:/home/jovyan/addresses.csv
```

Suppose we download the .csv file in our local root directory. And we use docker cp command to copy the file to the current container's working directory.



Now the .csv file had already under the jupyter notebook working directory.

Not secure | 172.31.10.126:20001/vnc_auto.html?token=0c1b7ce7-57d7-4247-a234-279db67a3fd9&instance_name=p930026096RoyMa5&instance_id=fa2af482-5fe6-4ae1-8eee-4021f8e88940

发送 Ctrl+Alt+Del 全部

Connected to p930026096RoyMa5 (instance.00001281)

Home x Untitled - Jupyter Notebook x Course: Data Processing Worl x +

0.0.0.0:8888/notebooks/Untitled.ipynb?kernel_name=python3

Jupyter Untitled (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

In [1]:

```
from pyspark.sql import SparkSession

spark = SparkSession.builder.appName("MyApp").getOrCreate()
df = spark.read.csv("addresses.csv")
df.show()
```

WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.spark.unsafe.Platform (file:/usr/local/spark-3.2.0-bin-hadoop3.2/jars/spark-unsafe-2.12-3.2.0.jar) to constructor java.nio.DirectByteBuffer(long,int)
WARNING: Please consider reporting this to the maintainers of org.apache.spark.unsafe.Platform
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
21/11/09 03:33:33 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-jav
a classes where applicable

| _c0 | _c1 | _c2 | _c3 | _c4 | _c5 |
|--------------------|----------|-------------------------|-------------|-----|-------|
| John | Doe | 120 jefferson st. | Riverside | NJ | 08075 |
| Jack | McGinnis | 220 hobo Av. | Phila | PA | 09119 |
| "John ""Da Man"" | Repici | 120 Jefferson St. | Riverside | NJ | 08075 |
| Stephen | Tyler | "7452 Terrace ""A..." | SomeTown | SD | 91234 |
| null | Blankman | null | SomeTown | SD | 00298 |
| "Joan ""the bone"" | Anne | Jet 9th, at Terrace plc | Desert City | CO | |

Create a python3 file. Type the sample code and run it. The result is shown on the above graph.