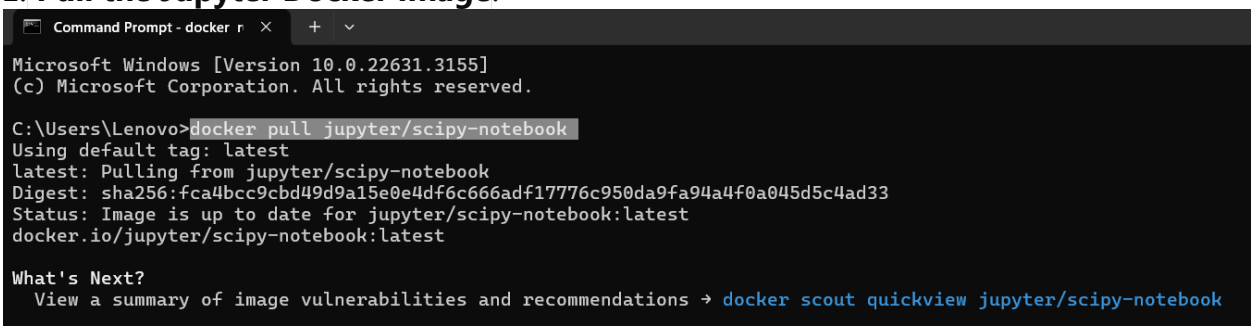
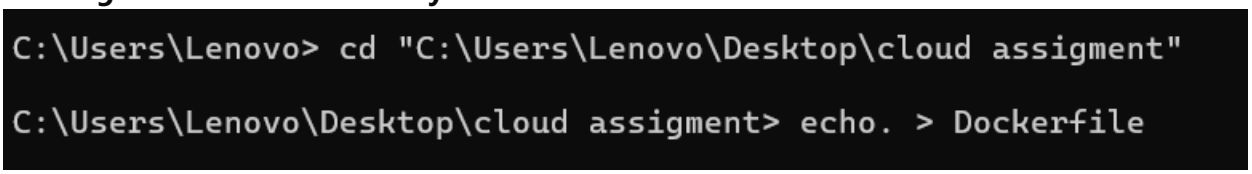
Name: habiba helal mohammed
AI LEVEL 2
ID:22010453
SECTION: WED 8.5

STEPS:
1.**Docker Setup** (DONE BY FOLLOWING STEPS IN SECTION 4)

2. **Pull the Jupyter Docker Image**:



3. **Navigate TO PROJECT directory and create docker file**



4. **Dockerfile Creation**

Dockerfile

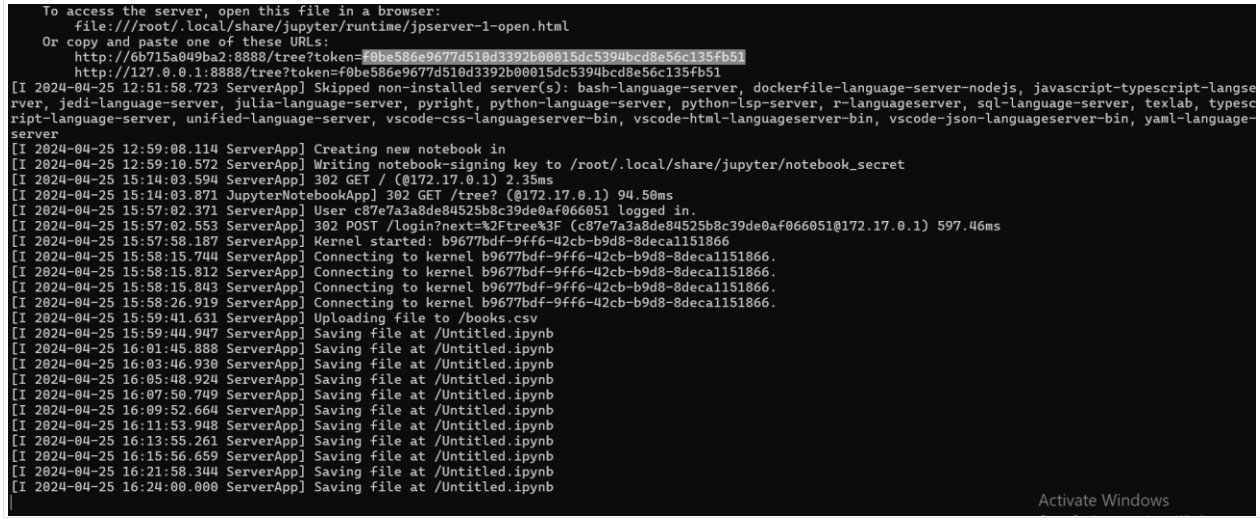
```
I: > Users > Lenovo > Desktop > cloud > Dockerfile > ...
1   # Use an official Python runtime as a parent image
2   FROM python:3.8
3
4   # Set the working directory to /app
5   WORKDIR /app
6
7   # Copy the current directory contents into the container at /app
8   COPY . /app
9
10  # Install any needed packages specified in requirements.txt
11  RUN pip install  --no-cache-dir -r requirements.txt
12
13  # Install Jupyter Notebook
14  RUN pip install jupyter
15
16  # Make port 8888 available to the world outside this container
17  EXPOSE 8888
18
19  # Define environment variable
20  ENV NAME World
21
22  # Run Jupyter Notebook when the container launches
23  CMD ["jupyter", "notebook", "--ip='0.0.0.0'", "--port=8888", "--no-browser", "--allow-root"]
24
```

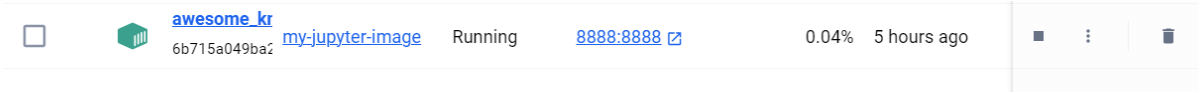
5. **Building the Docker Image**:

```
C:\Users\Lenovo\Desktop\cloud assigment> docker build -t my-jupyter-image .
[+] Building 0.0s (0/0)  docker:default
2024/04/25 14:45:15 http2: server: error reading preface from client //./pipe/docker_engine: file has already been close[+] Building 321.2s (9/9) FINISHED
                                                                                                                    docker:default
 => [internal] load build definition from Dockerfile                                                                          3.5s
 => => transferring dockerfile: 577B                                                                                          0.1s
 => [internal] load metadata for docker.io/library/python:3.8                                                                12.5s
 => [internal] load .dockerignore                                                                                             3.0s
 => => transferring context: 2B                                                                                               0.0s
 => [1/4] FROM docker.io/library/python:3.8@sha256:6ea16099cac9f66419d1fc3a63aaa9d783214e8e38d2a1c0db2bfb0852ef9b             0.0s
 => [internal] load build context                                                                                             1.4s
 => => transferring context: 577B                                                                                             0.0s
 => CACHED [2/4] WORKDIR /app                                                                                                 0.0s
 => [3/4] COPY . /app                                                                                                         3.3s
 => [4/4] RUN pip install jupyter                                                                                           269.0s
 => exporting to image                                                                                                       19.9s
 => => exporting layers                                                                                                      18.3s
 => => writing image sha256:808debff09df7d66d68f18cec7194013a4c7f097b1c25f5de90d345c4dd3559f                                  0.1s
 => => naming to docker.io/library/my-jupyter-image                                                                           0.2s

What's Next?
  View a summary of image vulnerabilities and recommendations → docker scout quickview
```

6. Run a Container from the Docker Image

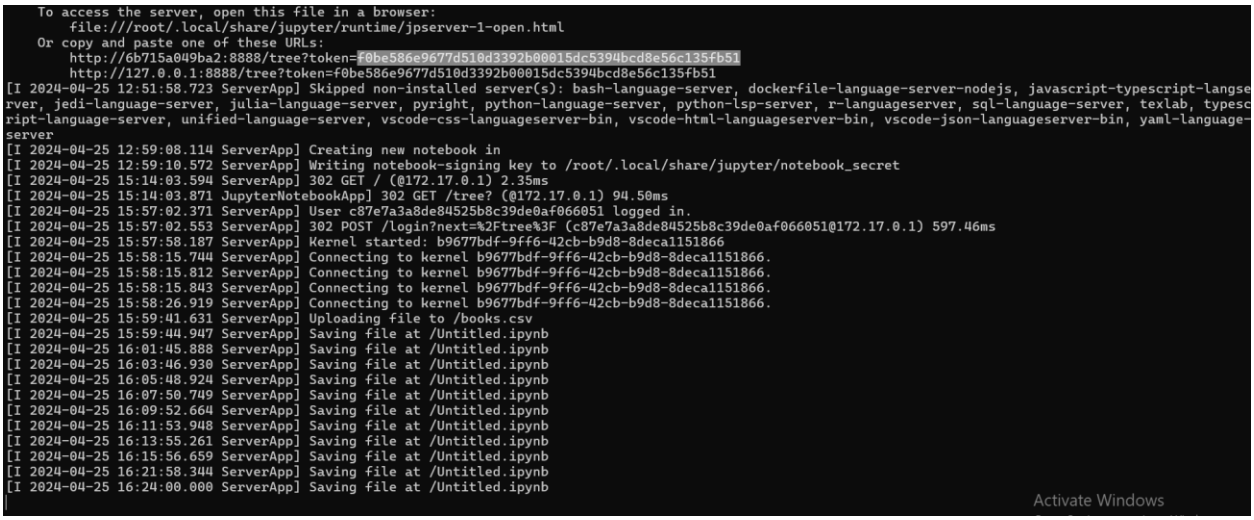




7. acess to jupyter server

`http://localhost:8888/`

`through this password`



jupyter

Password or token: •••••••••••••••••••••••••••••••••••• Log in

Token authentication is enabled

If no password has been configured, you need to open the 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 server list
```

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

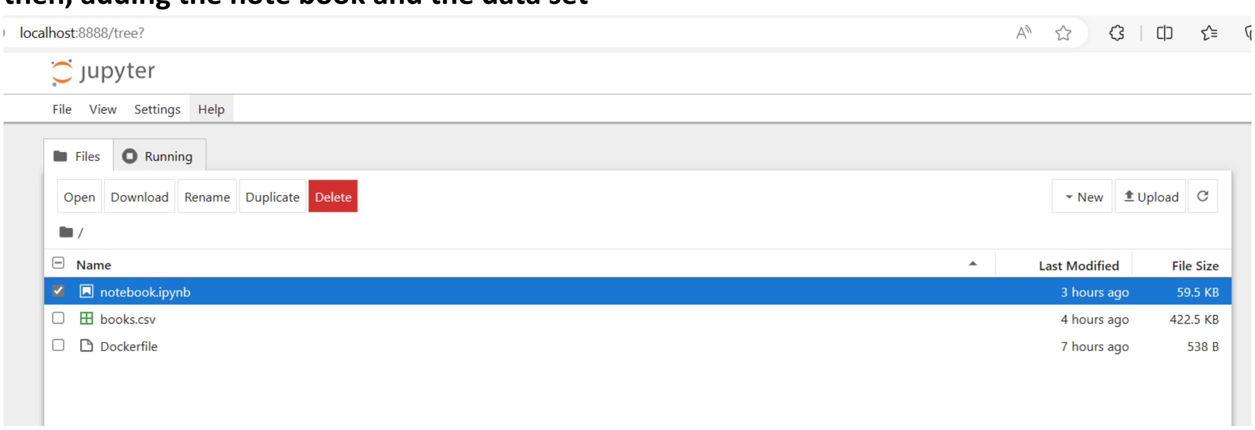
```
Currently running servers:
http://localhost:8888/?token=c8de56fa... :: /Users/you/notebooks
```

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

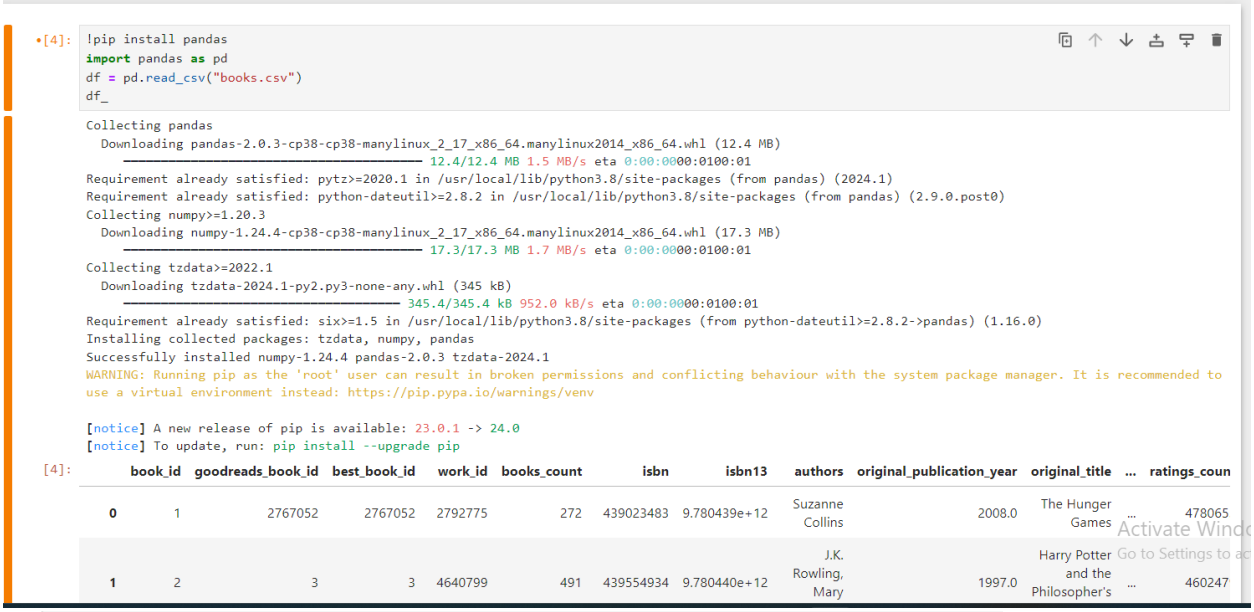
See **the documentation on how to enable a password** in place of token authentication, if you would like to avoid dealing with random tokens.

Cookies are required for authenticated access to the Jupyter server.

then, adding the note book and the data set



Finally , running the note book:



```
[4]: !pip install pandas
import pandas as pd
df = pd.read_csv("books.csv")
df_
```

```
Collecting pandas
  Downloading pandas-2.0.3-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (12.4 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.4/12.4 MB 1.5 MB/s eta 0:00:0000:0100:01
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.8/site-packages (from pandas) (2024.1)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.8/site-packages (from pandas) (2.9.0.post0)
Collecting numpy>=1.20.3
  Downloading numpy-1.24.4-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.3 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 17.3/17.3 MB 1.7 MB/s eta 0:00:0000:0100:01
Collecting tzdata>=2022.1
  Downloading tzdata-2024.1-py2.py3-none-any.whl (345 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 345.4/345.4 kB 952.0 kB/s eta 0:00:0000:0100:01
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.8/site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Installing collected packages: tzdata, numpy, pandas
Successfully installed numpy-1.24.4 pandas-2.0.3 tzdata-2024.1
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to
use a virtual environment instead: https://pip.pypa.io/warnings/venv

[notice] A new release of pip is available: 23.0.1 -> 24.0
[notice] To update, run: pip install --upgrade pip
```

[4]:

	book_id	goodreads_book_id	best_book_id	work_id	books_count	isbn	isbn13	authors	original_publication_year	original_title	...	ratings_coun
0	1	2767052	2767052	2792775	272	439023483	9.780439e+12	Suzanne Collins	2008.0	The Hunger Games	...	478065
1	2	3	3	4640799	491	439554934	9.780440e+12	J.K. Rowling, Mary	1997.0	Harry Potter and the Philosopher's	...	460247

]:
```
df.isnull().sum()
#df = df.dropna()
```

]:
```
book_id                        0
goodreads_book_id              0
best_book_id                   0
work_id                        0
books_count                    0
isbn                          52
isbn13                        44
authors                        0
original_publication_year      3
original_title                52
title                          0
language_code                109
average_rating                 0
ratings_count                  0
work_ratings_count             0
work_text_reviews_count        0
ratings_1                      0
ratings_2                      0
ratings_3                      0
ratings_4                      0
ratings_5                      0
image_url                      0
small_image_url                0
dtype: int64
```

]:
```
df = df.dropna()
df.drop_duplicates()
```

```
# Filter harry potter books
harry_potter_books = df[df['original_title'].str.contains("Harry Potter", case=False, na=False)]
harry_potter_books
```

	book_id	goodreads_book_id	best_book_id	work_id	books_count	isbn	isbn13	authors	original_publication_year	original_title	...	ratings_coun
1	2	3	3	4640799	491	439554934	9.780440e+12	J.K. Rowling, Mary GrandPré	1997.0	Harry Potter and the Philosopher's Stone	...	460247
6	18	5	5	2402163	376	043965548X	9.780440e+12	J.K. Rowling, Mary GrandPré, Rufus Beck	1999.0	Harry Potter and the Prisoner of Azkaban	...	183282
8	21	2	2	2809203	307	439358078	9.780439e+12	J.K. Rowling, Mary GrandPré	2003.0	Harry Potter and the Order of the Phoenix	...	173536
9	23	15881	15881	6231171	398	439064864	9.780439e+12	J.K. Rowling, Mary GrandPré	1998.0	Harry Potter and the Chamber of Secrets	...	177933
								J.K. Rowling		Harry Potter and the		

[14]:
```
# Find the most selling books
most_sold_books = harry_potter_books.sort_values(by='books_count', ascending=False)
print(most_sold_books[['original_title', 'books_count']])
```

```
                                       original_title  books_count
1                 Harry Potter and the Philosopher's Stone          491
9                  Harry Potter and the Chamber of Secrets          398
6                 Harry Potter and the Prisoner of Azkaban          376
10                     Harry Potter and the Goblet of Fire          332
8                Harry Potter and the Order of the Phoenix          307
12                  Harry Potter and the Half-Blood Prince          275
11                    Harry Potter and the Deathly Hallows          263
96                         Complete Harry Potter Boxed Set           76
1036  The Magical Worlds of Harry Potter: A Treasury...           42
613          Harry Potter Collection (Harry Potter, #1-6)            6
```

[15]:
```
# Calculate the average rating
average_rating = harry_potter_books['ratings_count'].mean()
print("Average Rating for Harry Potter Books:", average_rating)
```

```
Average Rating for Harry Potter Books: 1535692.9
```