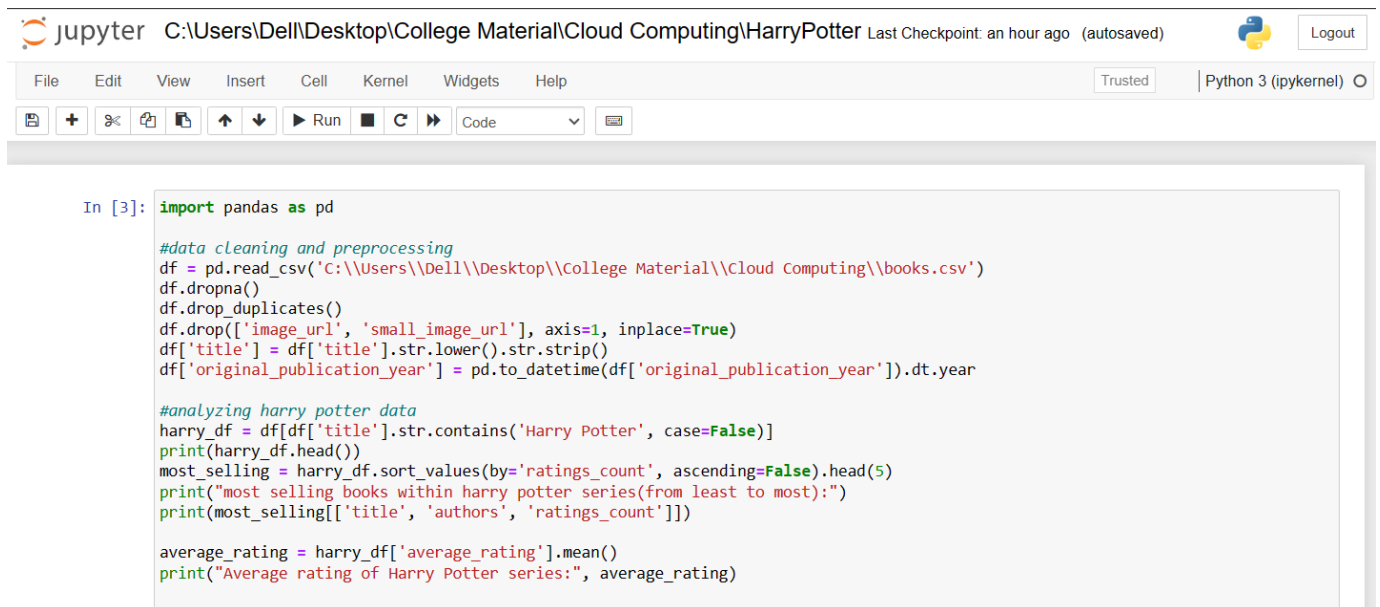


PULLED JUPYTER NOTEBOOK IMAGE:

```
C:\Users\Dell>docker images
REPOSITORY              TAG         IMAGE ID      CREATED        SIZE
jupyter/datascience-notebook  latest     f78a42f3bc9a  6 months ago  5.92GB
```

Wrote code on Jupyter notebook:



The screenshot shows a Jupyter Notebook interface. The top bar displays the Jupyter logo, the current file path 'C:\Users\Dell\Desktop\College Material\Cloud Computing\HarryPotter', and the last checkpoint status 'Last Checkpoint: an hour ago (autosaved)'. Below the top bar is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. A toolbar with various icons for file operations and execution is located below the menu bar. The main area of the notebook contains a code cell with the following Python code:

```
In [3]: import pandas as pd

#data cleaning and preprocessing
df = pd.read_csv('C:\\Users\\Dell\\Desktop\\College Material\\Cloud Computing\\books.csv')
df.dropna()
df.drop_duplicates()
df.drop(['image_url', 'small_image_url'], axis=1, inplace=True)
df['title'] = df['title'].str.lower().str.strip()
df['original_publication_year'] = pd.to_datetime(df['original_publication_year']).dt.year

#analyzing harry potter data
harry_df = df[df['title'].str.contains('Harry Potter', case=False)]
print(harry_df.head())
most_selling = harry_df.sort_values(by='ratings_count', ascending=False).head(5)
print("most selling books within harry potter series(from least to most):")
print(most_selling[['title', 'authors', 'ratings_count']])

average_rating = harry_df['average_rating'].mean()
print("Average rating of Harry Potter series:", average_rating)
```

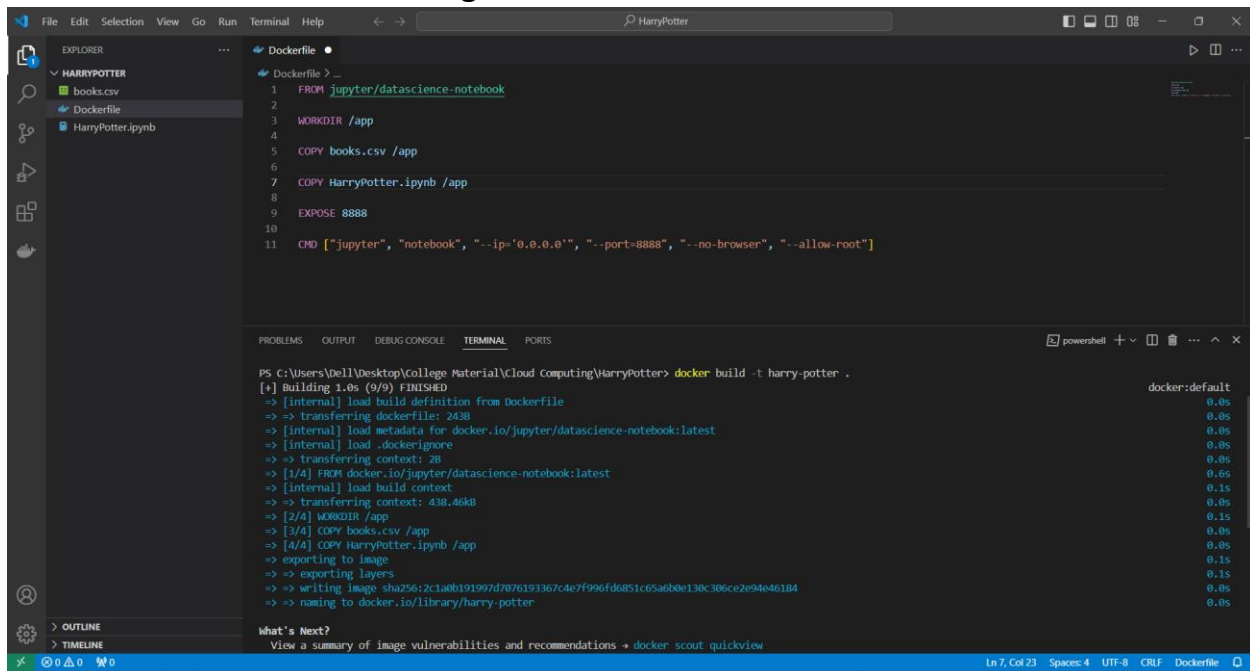
Output:

```
C:\Users\Dell\PycharmProjects\pythonProject5\venv\Scripts\python.exe C:\Users\Dell\PycharmProjects\pythonProject5\harrypteer.py
  book_id  goodreads_book_id  best_book_id  ...  ratings_3  ratings_4  ratings_5
1         2                 3          3  ...    455024    1156318    3011543
6        18                 5          5  ...    166129     509447    1266670
8        21                 2          2  ...    180210     494427    1124806
9        23            15881        15881  ...    242345     548266    1065084
10       24                 6          6  ...    151785     494926    1195045

[5 rows x 21 columns]
most selling books within harry potter series(from least to most):
      title  ... ratings_count
1  harry potter and the sorcerer's stone (harry p...  ...    4602479
6  harry potter and the prisoner of azkaban (harr...  ...    1832823
9  harry potter and the chamber of secrets (harry...  ...    1779331
10 harry potter and the goblet of fire (harry pot...  ...    1753043
11 harry potter and the deathly hallows (harry po...  ...    1746574

[5 rows x 3 columns]
Average rating of Harry Potter series: 4.482727272727273
```

Docker File and then built image:



The screenshot shows the PyCharm IDE interface. On the left, the Explorer pane shows a project named 'HARRYPOTTER' with files 'books.csv', 'Dockerfile', and 'HarryPotter.ipynb'. The main editor displays the content of the 'Dockerfile'.

```
Dockerfile
1 FROM jupyter/datascience-notebook
2
3 WORKDIR /app
4
5 COPY books.csv /app
6
7 COPY HarryPotter.ipynb /app
8
9 EXPOSE 8888
10
11 CMD ["jupyter", "notebook", "--ip='0.0.0.0'", "--port=8888", "--no-browser", "--allow-root"]
```

Below the editor, the 'TERMINAL' pane shows the output of the command `docker build -t harry-potter .`:

```
PS C:\Users\Dell\Desktop\College Material\Cloud Computing\HarryPotter> docker build -t harry-potter .
[*] Building 1.0s (9/9) FINISHED
-> [internal] load build definition from Dockerfile
-> => transferring dockerfile: 243B
-> [internal] load metadata for docker.io/jupyter/datascience-notebook:latest
-> [internal] load .dockerignore
-> => transferring context: 2B
-> [1/4] FROM docker.io/jupyter/datascience-notebook:latest
-> [internal] load build context
-> => transferring context: 438.4kB
-> [2/4] WORKDIR /app
-> [3/4] COPY books.csv /app
-> [4/4] COPY HarryPotter.ipynb /app
-> exporting to image
-> => exporting layers
-> writing image sha256:2c1a0b191997d7076193367c4e7f996fd0851c5a0b0e130c306ce2e94e6184
-> naming to docker.io/library/harry-potter
```

At the bottom of the terminal, there is a 'What's Next?' section with a link to 'View a summary of image vulnerabilities and recommendations -> docker scout quickview'.

Images:

3.84 GB / 5.92 GB in use

2 images

Last refresh: 9 hours ago

Search

<div></div>	Name	Tag	Status	Created	Size	Actions
<div></div>	<div><div>harry-potter</div><div>2c1a0b191997</div><div></div></div>	latest	<div>In use</div>	13 minutes ago	5.92 GB	<div><div></div><div></div><div></div></div>
<div></div>	<div><div>jupyter/datascience-notebook</div><div>f78a42f3bc9a</div><div></div></div>	latest	Unused	6 months ago	5.92 GB	<div><div></div><div></div><div></div></div>

Container:

Containers

Images

Volumes

Builds

Dev Environments BETA

Docker Scout

Extensions

Add Extensions

Search for images, containers, volumes, extensions... Ctrl+K

STATUS

Exited (0) (6 minutes ago)

Logs

Inspect

Bind mounts

Exec

Files

Stats

2024-04-23 22:45:04 [I 2024-04-23 20:45:04.257 ServerApp] nbdline | extension was successfully linked.

2024-04-23 22:45:04 [I 2024-04-23 20:45:04.264 ServerApp] notebook | extension was successfully linked.

2024-04-23 22:45:04 [I 2024-04-23 20:45:04.283 ServerApp] Writing Jupyter server cookie secret to /home/jovyan/.local/share/jupyter/runtime/jupyter_cookie_secret

2024-04-23 22:45:05 [I 2024-04-23 20:45:05.841 ServerApp] notebook_shim | extension was successfully linked.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.009 ServerApp] notebook_shim | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.016 ServerApp] jupyter_lsp | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.017 ServerApp] jupyter_server_mathjax | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.040 ServerApp] jupyter_server_proxy | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.042 ServerApp] jupyter_server_terminals | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.085 LabApp] JupyterLab extension loaded from /opt/conda/lib/python3.11/site-packages/jupyterlab

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.085 LabApp] JupyterLab application directory is /opt/conda/share/jupyter/lab

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.086 LabApp] Extension Manager is 'pypi'.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.091 ServerApp] jupyterlab | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.099 ServerApp] jupyterlab_git | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.116 ServerApp] nbclassic | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.345 ServerApp] nbdline | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.350 ServerApp] notebook | extension was successfully loaded.

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.353 ServerApp] Serving notebooks from local directory: /app

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.353 ServerApp] Jupyter Server 2.8.0 is running at:

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.353 ServerApp] http://4e7ee307e226:8888/tree?token=899f5a46aaa53dcb264325ce355d0e0df44dda64eea28689

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.353 ServerApp] http://127.0.0.1:8888/tree?token=899f5a46aaa53dcb264325ce355d0e0df44dda64eea28689

2024-04-23 22:45:06 [I 2024-04-23 20:45:06.353 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).

2024-04-23 22:45:06 [C 2024-04-23 20:45:06.360 ServerApp]

2024-04-23 22:45:06

2024-04-23 22:45:06 To access the server, open this file in a browser:

2024-04-23 22:45:06 file:///home/jovyan/.local/share/jupyter/runtime/jpservice-7-open.html

2024-04-23 22:45:06 Or copy and paste one of these URLs:

2024-04-23 22:45:06 http://4e7ee307e226:8888/tree?token=899f5a46aaa53dcb264325ce355d0e0df44dda64eea28689

2024-04-23 22:45:06 http://127.0.0.1:8888/tree?token=899f5a46aaa53dcb264325ce355d0e0df44dda64eea28689

2024-04-23 22:45:06 [I 2024-04-23 20:45:17.962 ServerApp] Skipped non-installed server(s): bash-language-server, dockerfile-language-server-nodejs, javascript-typescript-languageserver, jedi-language-server, julia-language-server, pyright, python-language-server, python-lsp-server, r-languageserver, sql-language-server, texlab, typescript-lan

Engine running

RAM 2.60 GB CPU 0.19% Signed in

New version available