

# Assignment Docker

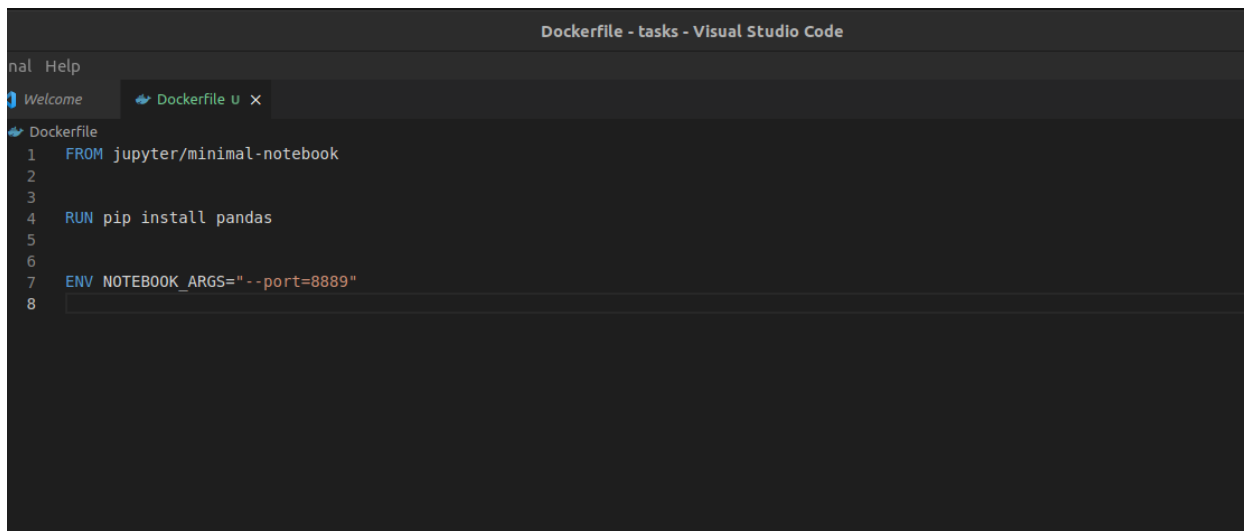
Completed by Muzammil Mehmood & Muhammad Talha  
& Maaz Javaid

## Problem Statement

- Build an image based on Jupyter Notebook (jupyter/minimal-notebook) with Pandas installed (pip install pandas)
- Create a container from this image and use the NOTEBOOK\_ARGS=--port=8889 environment variable to change the port Jupyter is exposed on
- Verify you can access it on port 8889 and that Pandas is installed (type import pandas in a notebook).

## Solution

Firstly we have created Dockerfile, to pull image of jupyter minimal notebook from dockerhub and install pandas into that container. The environment variable is NOTEBOOK\_ARGS=--port=8889.

A screenshot of a Visual Studio Code editor window titled "Dockerfile - tasks - Visual Studio Code". The editor shows a Dockerfile with the following content:

```
1 FROM jupyter/minimal-notebook
2
3
4 RUN pip install pandas
5
6
7 ENV NOTEBOOK_ARGS="--port=8889"
8
```

The interface includes a sidebar on the left with "Welcome" and "Dockerfile" tabs, and a top bar with "Dockerfile - tasks - Visual Studio Code".

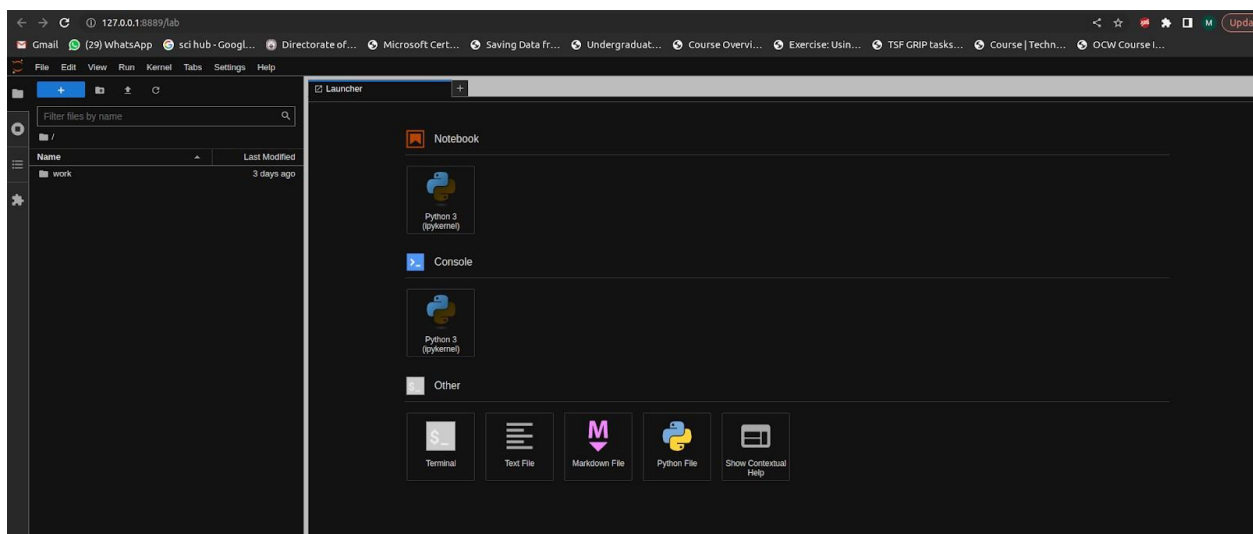
## Step 2 Building Docker Container

```
(bootcamp) muzammil@all-MS-7D35:~/GitRepo/data_engineering_bootcamp_2303/tasks$ docker build -t jupyter-pandas .
[+] Building 0.3s (6/6) FINISHED
=> [internal] load build definition from Dockerfile
=> => transferring dockerfile: 31B
=> [internal] load .dockerignore
=> => transferring context: 2B
=> [internal] load metadata for docker.io/jupyter/minimal-notebook:latest
=> [1/2] FROM docker.io/jupyter/minimal-notebook
=> CACHED [2/2] RUN pip install pandas
=> exporting to image
=> => exporting layers
=> writing image sha256:e0799b77b519214247367f27ece272d23084da8729550524ea402de9b3ad74aa
=> naming to docker.io/library/jupyter-pandas
(bootcamp) muzammil@all-MS-7D35:~/GitRepo/data_engineering_bootcamp_2303/tasks$
```

## Step 3 Running Docker Container

```
(bootcamp) muzammil@all-MS-7D35:~/GitRepo/data_engineering_bootcamp_2303/tasks$ docker run -p 8889:8889 jupyter-pandas
Entered start.sh with args: jupyter lab --port=8889
Executing the command: jupyter lab --port=8889
[I 2023-04-06 08:57:52.930 ServerApp] Package jupyterlab took 0.0000s to import
[I 2023-04-06 08:57:52.934 ServerApp] Package jupyter_server_fileid took 0.0035s to import
[I 2023-04-06 08:57:52.939 ServerApp] Package jupyter_server_terminals took 0.0044s to import
[I 2023-04-06 08:57:52.993 ServerApp] Package jupyter_server_ydoc took 0.0529s to import
[I 2023-04-06 08:57:52.995 ServerApp] Package nbclassic took 0.0000s to import
[W 2023-04-06 08:57:53.002 ServerApp] A `jupyter_server_extension_points` function was not found in nbclassic. Instead, a `_jupyter_server_extension_points` function will be deprecated in future releases of Jupyter Server.
[I 2023-04-06 08:57:53.004 ServerApp] Package notebook_shim took 0.0000s to import
[W 2023-04-06 08:57:53.004 ServerApp] A `jupyter_server_extension_points` function was not found in notebook_shim. Instead, a `_jupyter_server_extension_points` function will be deprecated in future releases of Jupyter Server.
[I 2023-04-06 08:57:53.014 ServerApp] jupyter_server_fileid | extension was successfully linked.
[I 2023-04-06 08:57:53.027 ServerApp] jupyter_server_terminals | extension was successfully linked.
[I 2023-04-06 08:57:53.039 ServerApp] jupyter_server_ydoc | extension was successfully linked.
```

## Step 4 Creating Jupyter Notebook



## Step 5 Importing Pandas as pd

