Data Science Learning Tools

ITCS 227 Intro to Data Science
Associate Prof. Dr. Suppawong Tuarob
Faculty of ICT, Mahidol University

Main Tools

- Language: Python
- IDE: Jupyter Notebook

Introduction to Python for Data Science

A Powerful Tool for Data Professionals

The Rise of Python in Data Science

- Most widely used programming language in data science.
- 2019 Kaggle Survey: ¾ of respondents use Python regularly.
- 2019 Glassdoor: >75% of data science job descriptions include Python.
- The preferred language for aspiring data scientists.

Key Features of Python

- High-level, general-purpose language.
- Large standard library with tools for various tasks:
  - Databases
  - Automation
  - Web scraping
  - Text processing
  - Image processing
  - Machine learning
  - Data analytics

Python's Data Science Toolkit

- Pandas: data manipulation
- NumPy: numerical computing
- SciPy: statistical computing
- Matplotlib: visualization

Introduction to Jupyter Notebooks

An Interactive Environment for Data Science

What is a Jupyter Notebook?

- Interactive Documents for Data Science
- Browser-based application for creating and sharing documents.
- Combines code, equations, visualizations, narrative text, links, etc.
- Analogy: A scientist's lab notebook for reproducible experiments.

Key Features of Jupyter Notebooks

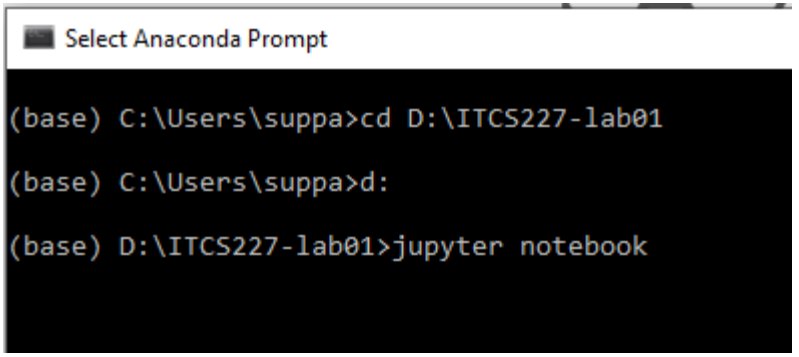
- Combines descriptive text, code blocks, and code output in a single file.
- Code execution generates output (plots, tables) within the notebook.
- Exportable to PDF or HTML for sharing.

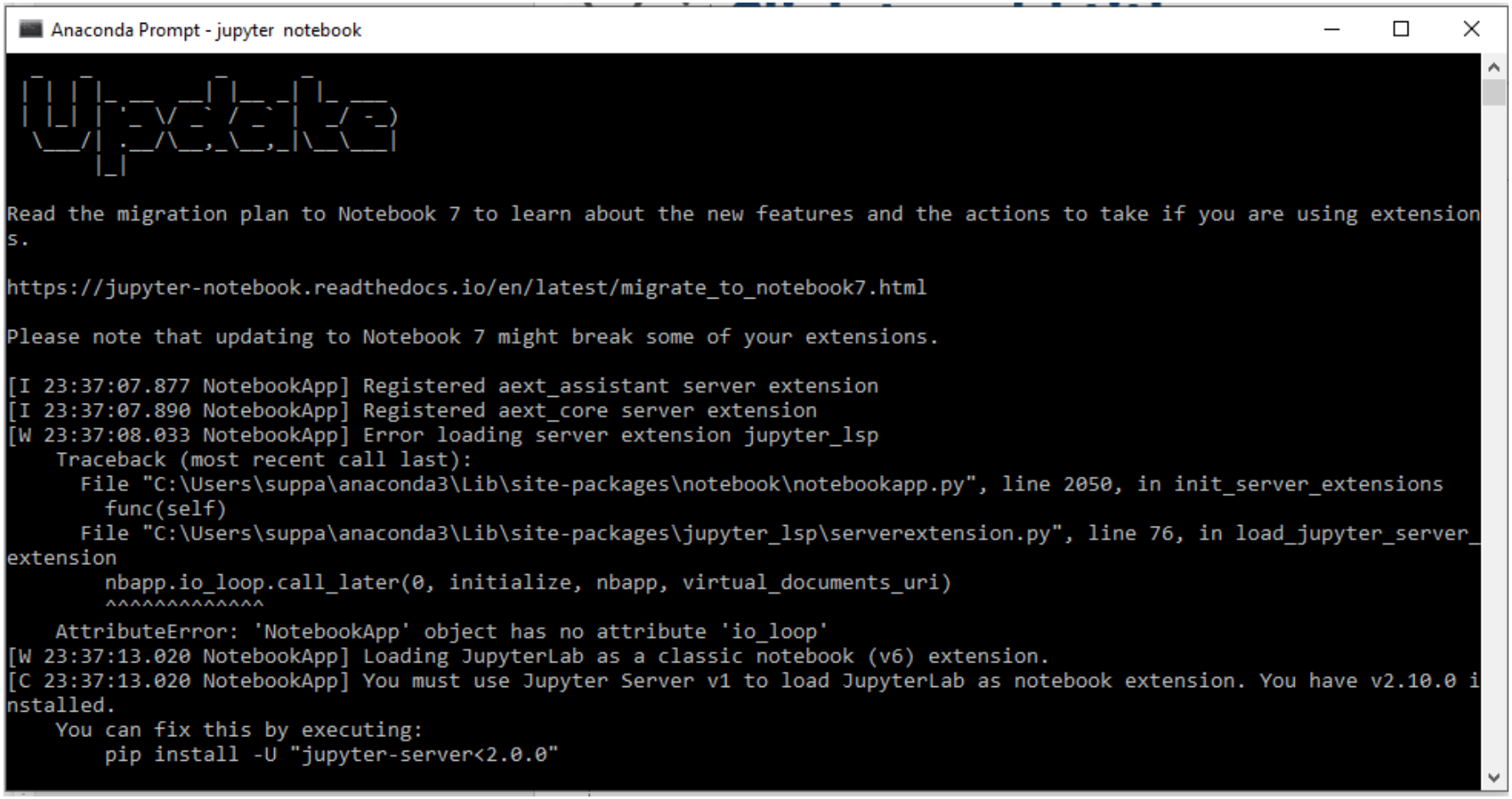
Getting Started with Jupyter Notebooks: Basic Operations

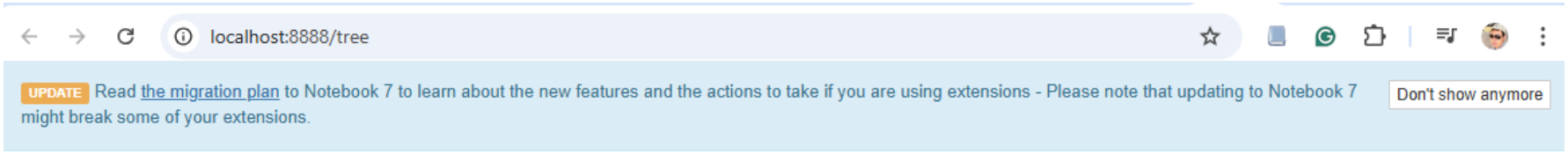
Working with Cells, Notebooks, and Presentations

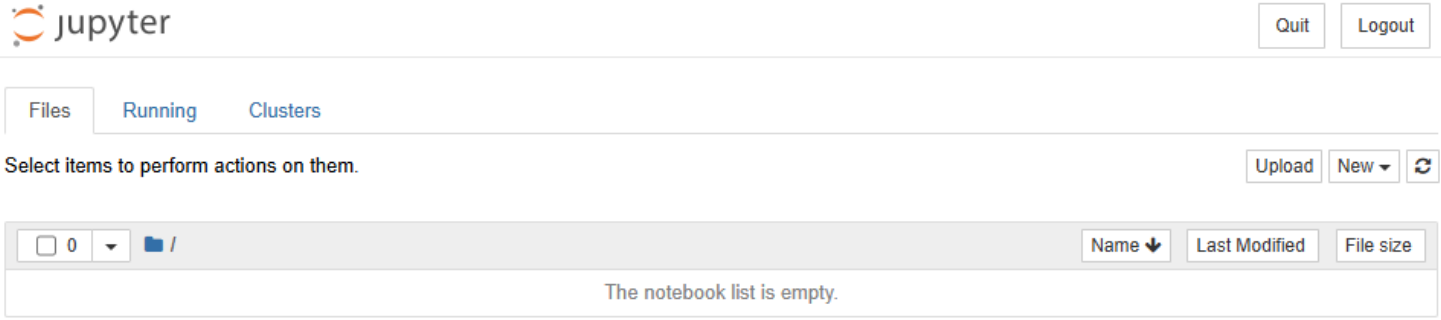
Accessing Jupyter for this Course

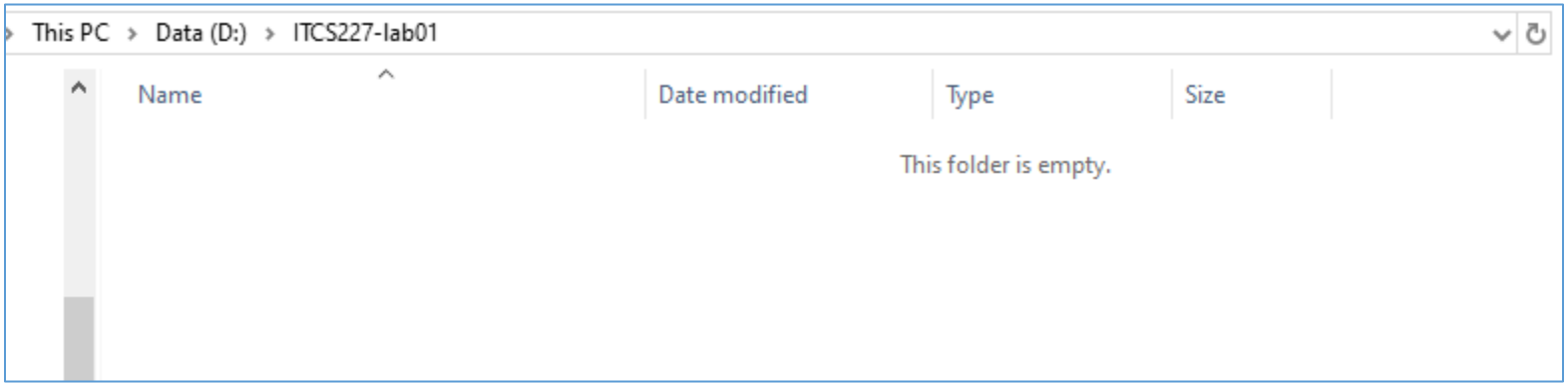
- Each lab computer should have Anaconda installed. If not, follow this tutorial: https://www.geeksforgeeks.org/install-jupyter-notebook-in-windows/
- Run “Anaconda Prompt”
- Navigate to your working folder, e.g., “D:\ITCS227-lab01”
- Run “jupyter notebook”

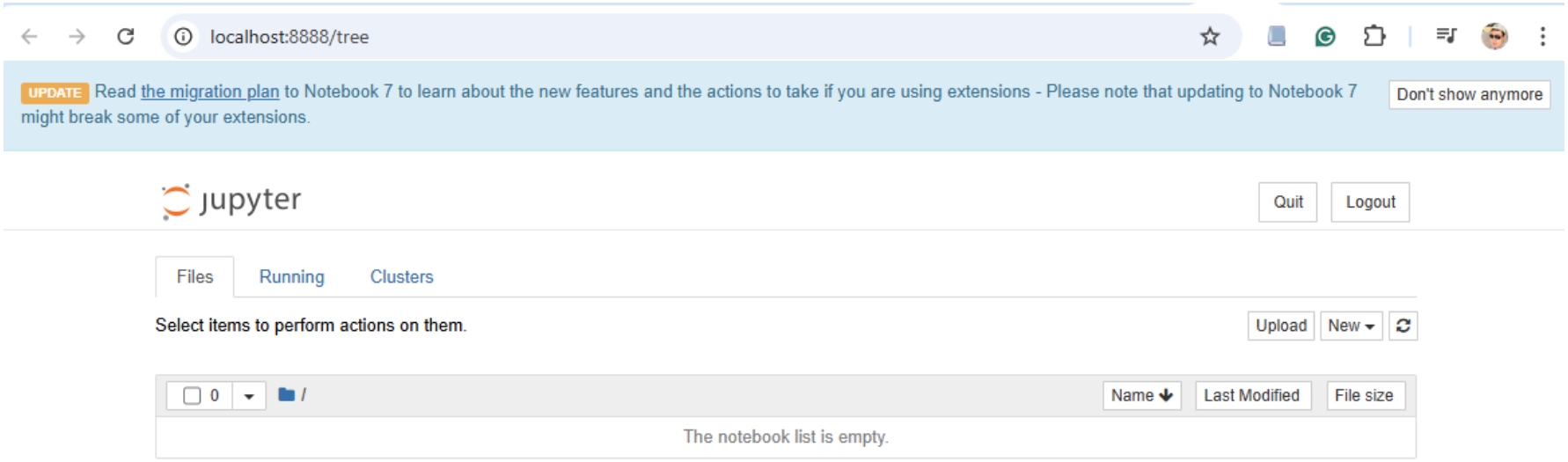


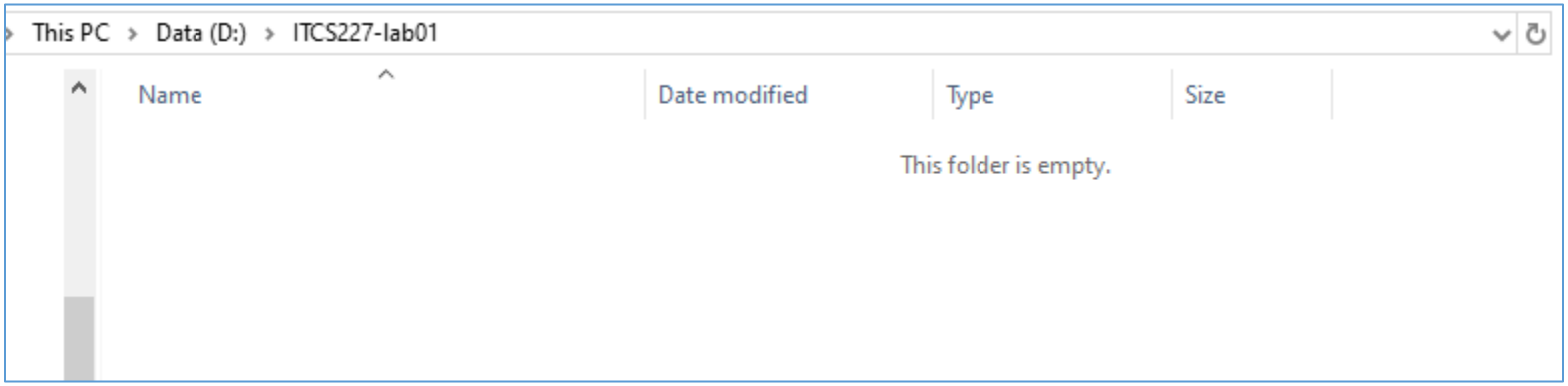

Anaconda Prompt - jupyter notebook
Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extension
s.
https://jupyter-notebook.readthedocs.io/en/latest/migrate_to_notebook7.html
Please note that updating to Notebook 7 might break some of your extensions.
[I 23:37:07.877 NotebookApp] Registered aext_assistant server extension
[I 23:37:07.890 NotebookApp] Registered aext_core server extension
[W 23:37:08.033 NotebookApp] Error loading server extension jupyter_lsp
Traceback (most recent call last):
File "C:\Users\suppa\anaconda3\Lib\site-packages\notebook\notebookapp.py", line 2050, in init_server_extensions
func(self)
File "C:\Users\suppa\anaconda3\Lib\site-packages\jupyter_lsp\serverextension.py", line 76, in load_jupyter_server_
extension
nbapp.io_loop.call_later(0, initialize, nbapp, virtual_documents_uri)
^^^^^^^^^^^^^
AttributeError: 'NotebookApp' object has no attribute 'io_loop'
[W 23:37:13.020 NotebookApp] Loading JupyterLab as a classic notebook (v6) extension.
[C 23:37:13.020 NotebookApp] You must use Jupyter Server v1 to load JupyterLab as notebook extension. You have v2.10.0 i
nstalled.
You can fix this by executing:
pip install -U "jupyter-server<2.0.0"

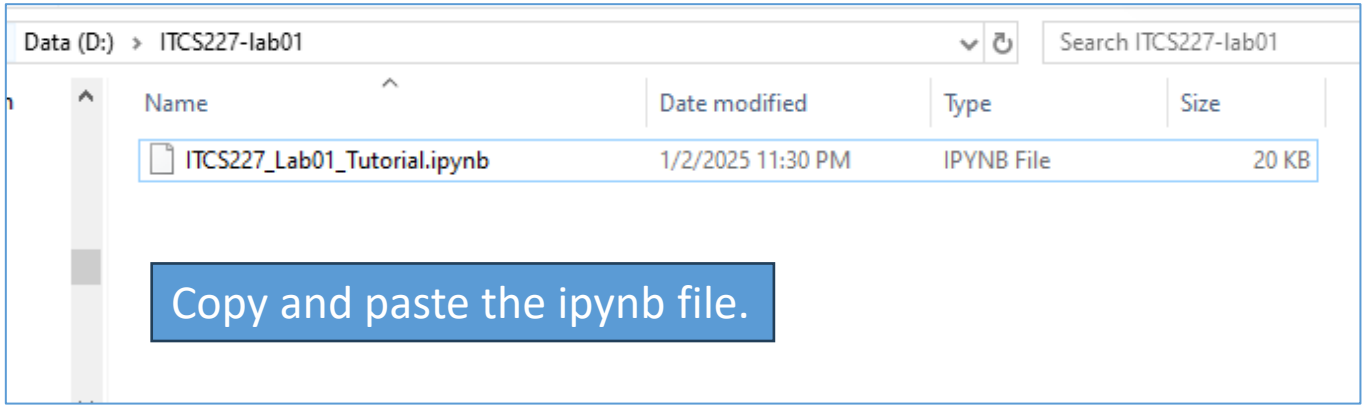

localhost:8888/tree
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore

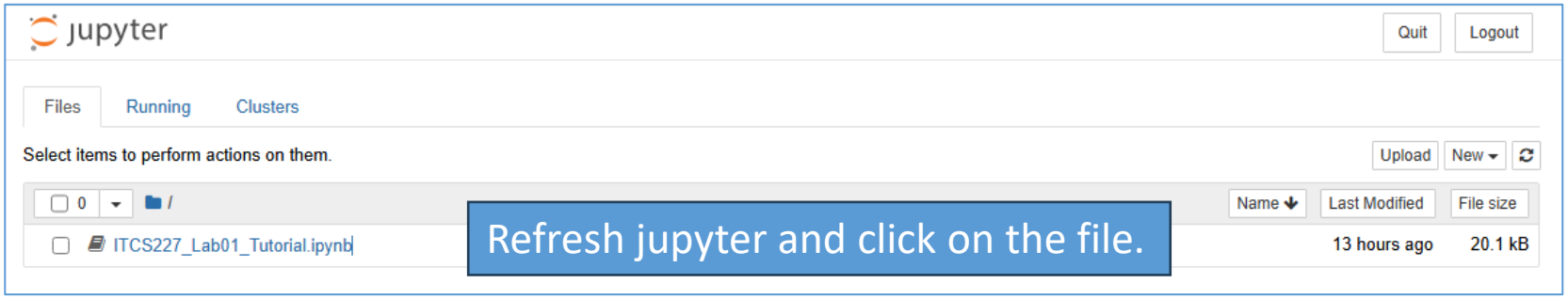

jupyter
Quit
Logout
Files
Running
Clusters
Select items to perform actions on them.
Upload
New
0
/
Name
Last Modified
File size
The notebook list is empty.

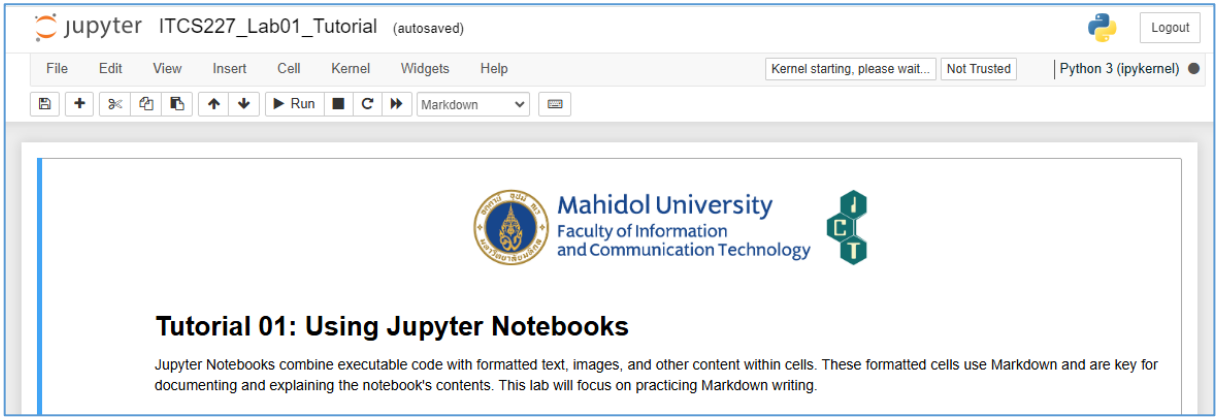

This PC
Data (D:)
ITCS227-lab01
Name
Date modified
Type
Size
This folder is empty.


localhost:8888/tree
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter
Quit
Logout
Files
Running
Clusters
Select items to perform actions on them.
Upload
New
0
/
Name
Last Modified
File size
The notebook list is empty.


This PC
Data (D:)
ITCS227-lab01
Name
Date modified
Type
Size
This folder is empty.


Data (D:) › ITCS227-lab01
Search ITCS227-lab01
Name
Date modified
Type
Size
ITCS227_Lab01_Tutorial.ipynb
1/2/2025 11:30 PM
IPYNB File
20 KB
Copy and paste the ipynb file.


jupyter
Quit
Logout
Files
Running
Clusters
Select items to perform actions on them.
Upload
New
0
/
Name
Last Modified
File size
ITCS227_Lab01_Tutorial.ipynb
13 hours ago
20.1 kB
Refresh jupyter and click on the file.


jupyter ITCS227_Lab01_Tutorial (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Kernel starting, please wait...
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Mahidol University
Faculty of Information
and Communication Technology
Tutorial 01: Using Jupyter Notebooks
Jupyter Notebooks combine executable code with formatted text, images, and other content within cells. These formatted cells use Markdown and are key for documenting and explaining the notebook's contents. This lab will focus on practicing Markdown writing.

Let’s go to the lab