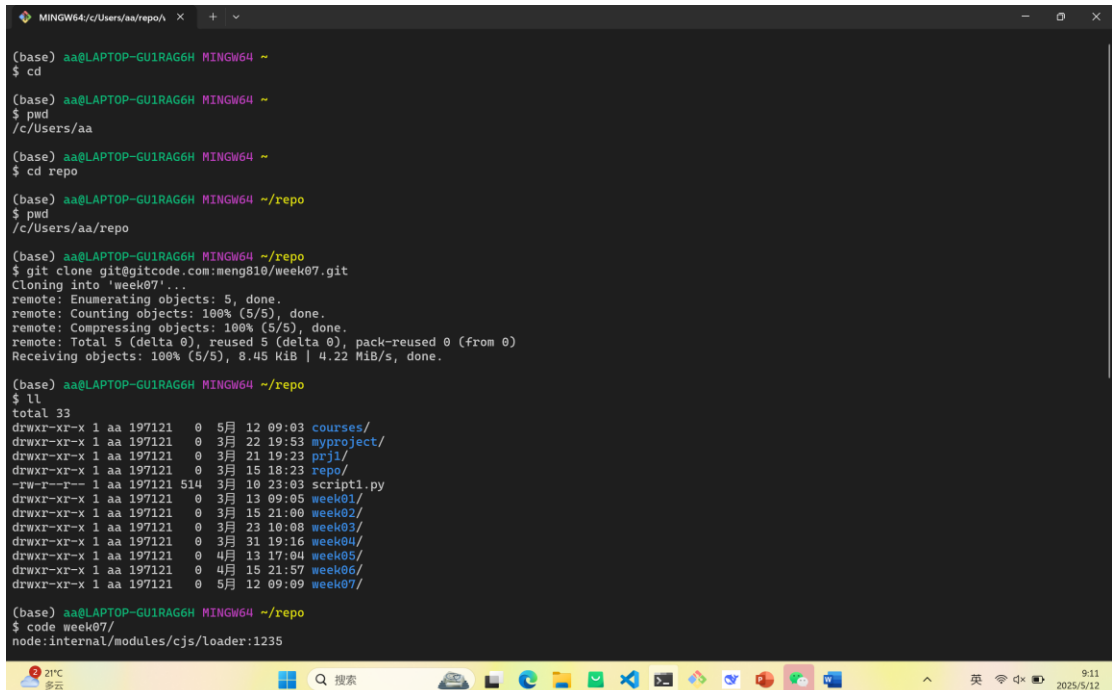


1. Fork 第 07 周打卡 仓库至你的名下，然后将你名下的这个仓库 Clone 到你的本地计算机



```
(base) aa@LAPTOP-GUIRAG6H MINGW64 ~
$ cd

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~
$ pwd
/c/Users/aa

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~
$ cd repo

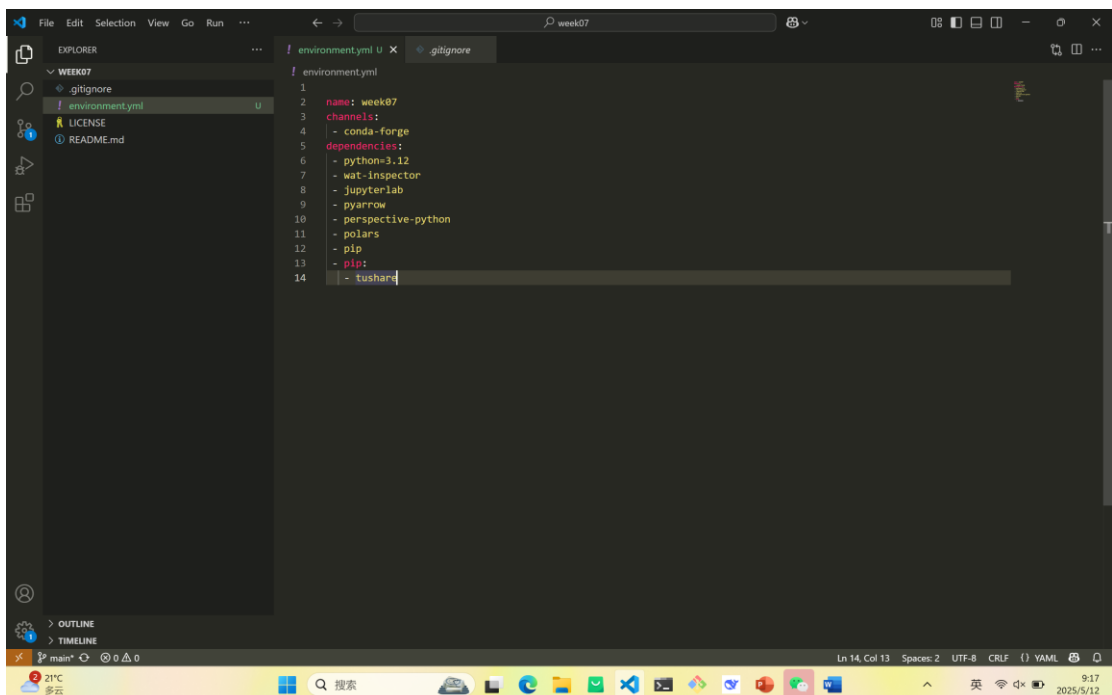
(base) aa@LAPTOP-GUIRAG6H MINGW64 ~/repo
$ pwd
/c/Users/aa/repo

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~/repo
$ git clone git@github.com:meng810/week07.git
Cloning into 'week07'...
remote: Enumerating objects: 5, done.
remote: Counting objects: 100% (5/5), done.
remote: Compressing objects: 100% (5/5), done.
remote: Total 5 (delta 0), reused 5 (delta 0), pack-reused 0 (from 0)
Receiving objects: 100% (5/5), 8.45 KiB | 4.22 MiB/s, done.

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~/repo
$ ll
total 33
drwxr-xr-x 1 aa 197121 0 5月 12 09:03 courses/
drwxr-xr-x 1 aa 197121 0 3月 22 19:53 myproject/
drwxr-xr-x 1 aa 197121 0 3月 21 19:23 prj1/
drwxr-xr-x 1 aa 197121 0 3月 15 18:23 repo/
-rw-r--r-- 1 aa 197121 514 3月 10 23:03 script1.py
drwxr-xr-x 1 aa 197121 0 3月 13 09:05 week01/
drwxr-xr-x 1 aa 197121 0 3月 15 21:00 week02/
drwxr-xr-x 1 aa 197121 0 3月 23 10:08 week03/
drwxr-xr-x 1 aa 197121 0 3月 31 19:16 week04/
drwxr-xr-x 1 aa 197121 0 4月 13 17:04 week05/
drwxr-xr-x 1 aa 197121 0 4月 15 21:57 week06/
drwxr-xr-x 1 aa 197121 0 5月 12 09:09 week07/

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~/repo
$ code week07/
node:internal/modules/cjs/loader:1235
```

2. 用 VS Code 打开项目目录，新建一个 environment.yml 文件，指定安装 Python 3.12 和 jupyterlab，然后运行 conda env create 命令创建 Conda 环境



```
File Edit Selection View Go Run ...
week07
.gitignore
environment.yml
LICENSE
README.md

environment.yml
1
2 name: week07
3 channels:
4 - conda-forge
5 dependencies:
6 - python=3.12
7 - wat-inspector
8 - jupyterlab
9 - pyarrow
10 - perspective-python
11 - polars
12 - pip
13 - pip:
14 - tushare
```

```
MINGW64/c:/Users/aa/repo/ X + v
Downloading https://mirrors.tuna.tsinghua.edu.cn/pypi/web/packages/38/33/0b4ed09bfff3b69887c2545f1c587420c89685d36c377095e465cc00759b/tushare-1.4)
Requirement already satisfied: pandas in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: requests in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Collecting lxml (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7))
Downloading https://mirrors.tuna.tsinghua.edu.cn/pypi/web/packages/91/1e/05ddcb57ad2f3069101611bd5f5084157d90861a2ef460bf42f45ccdd944/lxml-5.4.0-0)
3.8/3.8 MB 6.7 MB/s eta 0:00:00
Collecting simplejson (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7))
Downloading https://mirrors.tuna.tsinghua.edu.cn/pypi/web/packages/bc/5d/4e243e9377a3560107c69f6f7c2eed8589163f5ed14324e864871daa2dd9/simplejson-0)
Collecting bs4 (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7))
Using cached https://mirrors.tuna.tsinghua.edu.cn/pypi/web/packages/51/bb/bf772a159614954d84aa832c129624ba6c32faa559dfb200a534e50b/bs4-0.0.2-0)
Requirement already satisfied: websocket-client>=0.57.0 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Collecting tqdm (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7\site-packages (from tushare->r C:\Users\aa\repo\week07\condaenv.vw7))
Downloading https://mirrors.tuna.tsinghua.edu.cn/pypi/web/packages/d0/30/dc54f88dd4a2b5dc8a0279bdd7270e735851848b762aeb1c1184ed1f6b14/tqdm-4.67.1)
Requirement already satisfied: beautifulsoup4 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from bs4->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: soupsieve>1.2 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from beautifulsoup4->bs4->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: typing-extensions>=4.0.0 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from beautifulsoup4->bs4->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: numpy>=1.26.0 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from pandas->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from pandas->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: pytz>=2020.1 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from pandas->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: tzdata>=2022.7 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from pandas->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: six>=1.5 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from python-dateutil>=2.8.2->pandas->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from requests->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: idna<4,>=2.5 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from requests->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from requests->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from requests->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Requirement already satisfied: colorama in c:\users\aa\anaconda3\envs\week07\lib\site-packages (from requests->tushare->r C:\Users\aa\repo\week07\condaenv.vw7)
Installing collected packages: tqdm, simplejson, lxml, bs4, tushare
Successfully installed bs4-0.0.2 lxml-5.4.0 simplejson-3.20.1 tqdm-4.67.1 tushare-1.4.21

done
#
# To activate this environment, use
#
# $ conda activate week07
#
# To deactivate an active environment, use
#
# $ conda deactivate
#

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~/repo/week07 (main)
$
```

3. 在项目目录下，运行 jupyter lab 命令，启动 后端 (Backend) 服务，在浏览器里粘贴地址访问 前端 (Frontend) 页面

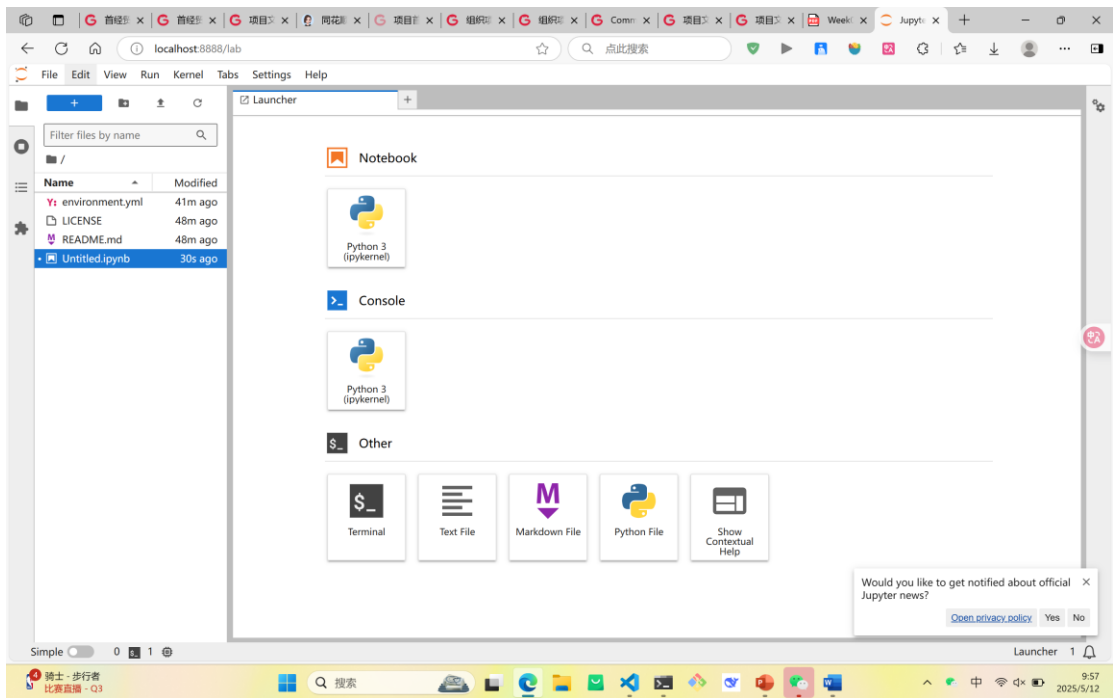
后端 (Backend) 服务是指在计算机系统或网络架构中，位于服务器端，为前端应用程序提供支持和数据处理的一系列功能和服务的统称。它主要负责处理业务逻辑、数据存储与管理、系统间交互等任务，不直接用户进行交互，却对整个系统的稳定运行和功能实现起着关键作用。

前端 (Frontend) 在计算机系统或应用程序中，是用户直接与之交互的部分，它负责将数据和功能以可视化、可操作的方式呈现给用户，主要涉及用户界面 (UI) 设计和用户体验 (UX) 优化，旨在让用户能够轻松、高效地使用系统或应用程序。

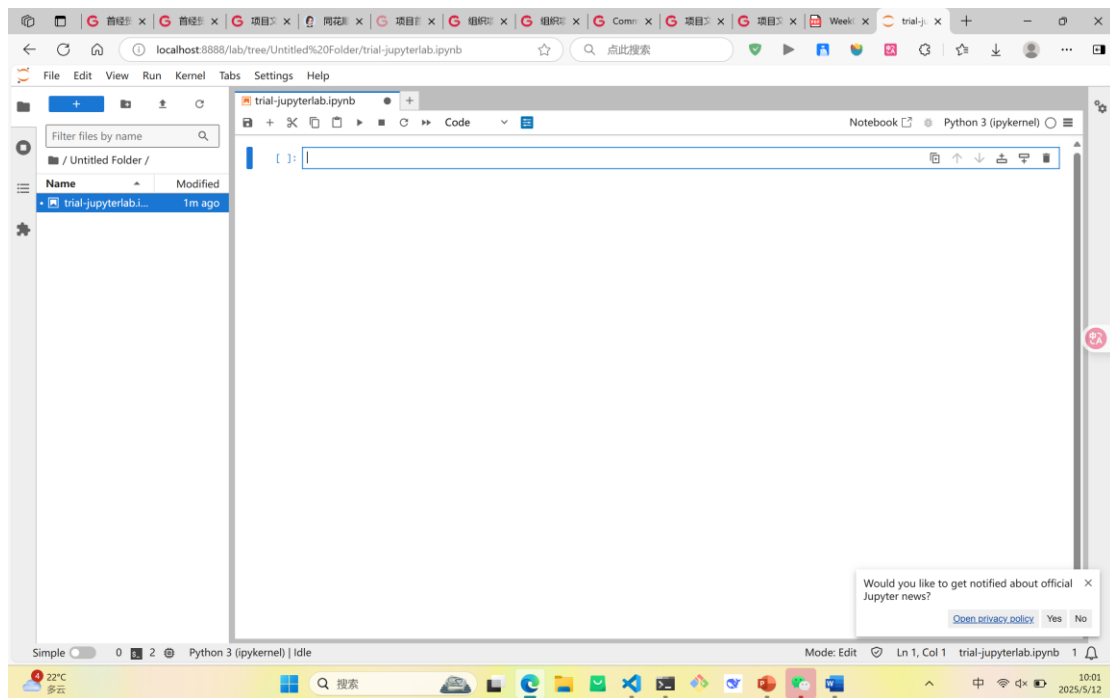
```
MINGW64/c:/Users/aa/repo/ X + v
#
# To deactivate an active environment, use
#
# $ conda deactivate
#

(base) aa@LAPTOP-GUIRAG6H MINGW64 ~/repo/week07 (main)
$ jupyter lab
[2025-05-12 09:54:14.927 ServerApp] jupyter_lsp | extension was successfully linked.
[2025-05-12 09:54:14.934 ServerApp] jupyter_server_terminals | extension was successfully linked.
[2025-05-12 09:54:14.943 ServerApp] jupyterlab | extension was successfully linked.
[2025-05-12 09:54:14.957 ServerApp] notebook | extension was successfully linked.
[2025-05-12 09:54:14.964 ServerApp] Writing Jupyter server cookie secret to C:\Users\aa\AppData\Roaming\jupyter\runtime\jupyter_cookie_secret
[2025-05-12 09:54:15.774 ServerApp] notebook_shim | extension was successfully linked.
[2025-05-12 09:54:15.774 ServerApp] panel.io.jupyter_server_extension | extension was successfully linked.
[2025-05-12 09:54:15.874 ServerApp] notebook_shim | extension was successfully loaded.
[2025-05-12 09:54:15.876 ServerApp] jupyter_lsp | extension was successfully loaded.
[2025-05-12 09:54:15.878 ServerApp] jupyter_server_terminals | extension was successfully loaded.
[2025-05-12 09:54:15.881 LabApp] JupyterLab extension loaded from C:\Users\aa\anaconda3\lib\site-packages\jupyterlab
[2025-05-12 09:54:15.881 LabApp] JupyterLab application directory is C:\Users\aa\anaconda3\share\jupyter\lab
[2025-05-12 09:54:15.889 LabApp] Extension Manager is 'pypi'.
[2025-05-12 09:54:16.174 ServerApp] jupyterlab | extension was successfully loaded.
[2025-05-12 09:54:16.183 ServerApp] notebook | extension was successfully loaded.
[2025-05-12 09:54:16.183 ServerApp] panel.io.jupyter_server_extension | extension was successfully loaded.
[2025-05-12 09:54:16.185 ServerApp] Serving notebooks from local directory: C:\Users\aa\repo\week07
[2025-05-12 09:54:16.185 ServerApp] Jupyter Server 2.14.1 is running at:
[2025-05-12 09:54:16.185 ServerApp] http://localhost:8888/lab?token=cfb778a26d4a4819178a77a41a4e93ff309acce15103f33a
[2025-05-12 09:54:16.187 ServerApp] http://127.0.0.1:8888/lab?token=cfb778a26d4a4819178a77a41a4e93ff309acce15103f33a
[2025-05-12 09:54:16.187 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 2025-05-12 09:54:16.276 ServerApp]

To access the server, open this file in a browser:
file:///C:/Users/aa/AppData/Roaming/jupyter/runtime/jpserver-28028-open.html
Or copy and paste one of these URLs:
http://localhost:8888/lab?token=cfb778a26d4a4819178a77a41a4e93ff309acce15103f33a
http://127.0.0.1:8888/lab?token=cfb778a26d4a4819178a77a41a4e93ff309acce15103f33a
[2025-05-12 09:54:16.276 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, r-languageserver, sql-language-server, texlab, typescript-language-server, unified-languageserver, vscode-css-languageserver-bin, vscode-html-languageserver-bin, vscode-json-languageserver-bin, yamllang-language-server
0.00s - Debugger warning: It seems that frozen modules are being used, which may
0.00s - make the debugger miss breakpoints. Please pass -Xfrozen_modules=off
0.00s - to python to disable frozen modules.
```



4. 在 JupyterLab 页面里，新建一个 Notebook，改名为 trial-jupyterlab.ipynb，在里面实践掌握以下功能：



在单元格 (Cell) 里编写 Python 代码，按 Shift+Enter 运行 Cell 并下移

```
[1]: x=5
      x=3*5
      print(x)
```

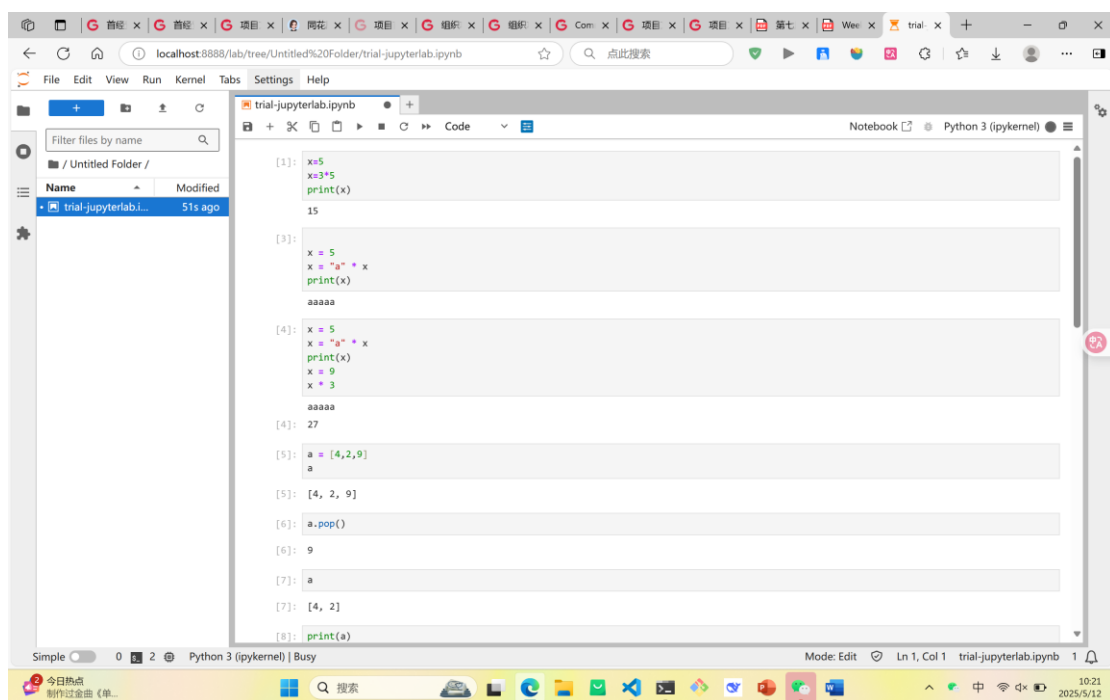
15

在单元格 (Cell) 上按 ESC 切换到 **命令模式** (command mode), 按 Enter 切换到 **编写模式** (edit mode)

单元格 (Cell) 序号为 * 表示代码运行中, 尚未返回, 按 ii 可以打断 (KeyboardInterrupt) (类似于终端的 Ctrl+C)

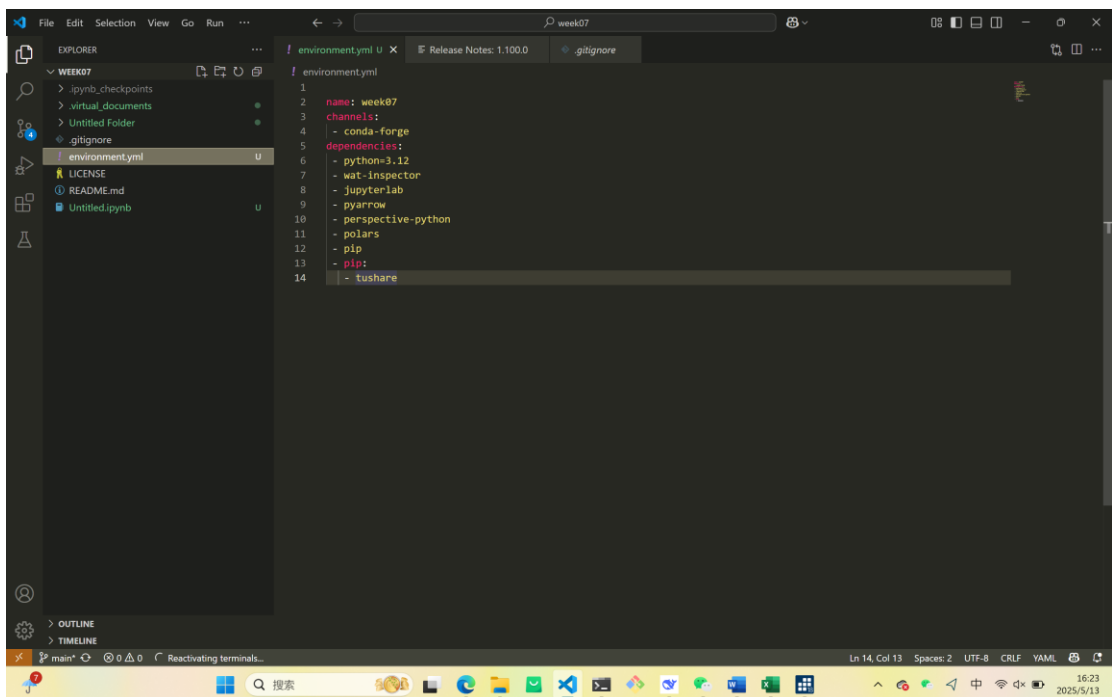
在单元格 (Cell) 的命令模式下, 按 00 重启后端 Python 解释器 (被 Jupyter 称为 Kernel), 重启后需要从上至下重新运行一遍代码 (Shift+Enter), 运行前建议先在菜单里选择 “Edit / Clear Outputs of All Cells” 清空全部页面显示的输出

在单元格 (Cell) 的命令模式下, 按 m 切换至 Markdown 模式, 按 y 切换至 Python 模式



关闭前端页面, 在后端按 Ctrl+C 打断运行中的服务, 回到 Bash 提示符

5. 通过 tushare 软件包下载保存一些数据
修改 environment.yml 文件, 添加 pip: tushare



在终端 (Terminal) 激活 week07 Conda 环境, 运行 ipython 命令启动 IPython 交互界面 ([IPython](#) 是 Jupyter 项目的一部份, ipython 是 jupyterlab 的依赖项之一)

```
(base) aa@LAPTOP-GU1RAG6H MINGW64 ~/repo/week07 (main)
$ conda activate week07
(week07)
aa@LAPTOP-GU1RAG6H MINGW64 ~/repo/week07 (main)
$ conda deactivate
(base)
aa@LAPTOP-GU1RAG6H MINGW64 ~/repo/week07 (main)
```

```
IPython: C:\repo\week07
widgetsnbextension 3.6.6 py312haa95532_0 https://repo.anaconda.com/pkgs/main
win_inet_pton 1.1.0 py312haa95532_0 https://repo.anaconda.com/pkgs/main
winpty 0.4.3 4 https://repo.anaconda.com/pkgs/main
wrapt 1.14.1 py312h2bbff1b_0 https://repo.anaconda.com/pkgs/main
xarray 2023.6.0 py312haa95532_0 https://repo.anaconda.com/pkgs/main
xlwings 0.32.1 py312haa95532_0 https://repo.anaconda.com/pkgs/main
xyzservices 2022.9.0 py312haa95532_1 https://repo.anaconda.com/pkgs/main
xz 5.4.6 h8ce25b3_1 https://repo.anaconda.com/pkgs/main
yaml 0.2.5 he774522_0 https://repo.anaconda.com/pkgs/main
yaml-cpp 0.8.0 hd77b12b_1 https://repo.anaconda.com/pkgs/main
yapf 0.40.2 py312haa95532_0 https://repo.anaconda.com/pkgs/main
yarl 1.11.0 py312h827c3e9_0 https://repo.anaconda.com/pkgs/main
zeromq 4.3.5 hd77b12b_0 https://repo.anaconda.com/pkgs/main
zfp 1.0.0 hd77b12b_0 https://repo.anaconda.com/pkgs/main
zict 3.0.0 py312haa95532_0 https://repo.anaconda.com/pkgs/main
zipp 3.17.0 py312haa95532_0 https://repo.anaconda.com/pkgs/main
zlib 1.2.13 h8ce25b3_1 https://repo.anaconda.com/pkgs/main
zlib-ng 2.0.7 h2bbff1b_0 https://repo.anaconda.com/pkgs/main
zope 1.0 py312haa95532_1 https://repo.anaconda.com/pkgs/main
zope.interface 5.4.0 py312h2bbff1b_0 https://repo.anaconda.com/pkgs/main
zstandard 0.23.0 py312h4fc1ca9_0 https://repo.anaconda.com/pkgs/main
zstd 1.5.6 h8880b57_0 https://repo.anaconda.com/pkgs/main
(base)
aa8LAPTOP-GUIRAG6H MINGW64 ~/repo/week07 (main)
$ ipython
Python 3.12.7 | packaged by Anaconda, Inc. | (main, Oct 4 2024, 13:17:27) [MSC v.1929 64 bit (AMD64)]
Type 'copyright', 'credits' or 'license' for more information
IPython 8.27.0 -- An enhanced Interactive Python. Type '?' for help.

In [1]:
```

6.Parquet 格式和 CSV 格式的特点和适用领域:

Parquet

特点:

- **列式存储**: 按列组织数据, 适合只查询部分列的场景 (如 **SELECT column1 FROM table**)。
- **高效压缩**: 支持 Snappy、Gzip 等压缩算法, 显著减少存储成本。
- **复杂类型支持**: 可存储嵌套数据 (如 JSON、数组、Map)。
- **谓词下推**: 过滤条件下推到存储层, 减少 I/O (如 **WHERE age > 30** 只读取相关数据块)。

适用场景:

- **大数据分析** (Spark、Hive、BigQuery 等)
- **数据仓库/湖仓一体** (Delta Lake、Iceberg)
- **机器学习特征存储** (需快速读取部分列)

CSV

特点:

- **人类可读**: 纯文本格式, 可直接用文本编辑器查看/编辑。
- **简单通用**: 所有编程语言和工具均支持。

- **无模式约束**：灵活但易出错（如列类型不一致）。

适用场景：

- **数据交换**（导出给非技术人员）
- **临时存储/小数据集**（<1GB）
- **传统系统接口**（银行/ERP 系统导出）

7.Bar Chart 的概念、适用情形以及实现代码

概念

条形图是一种以长方形的长度为变量的统计图表。在条形图中，横轴（也称为 x 轴）表示不同的类别，纵轴（也称为 y 轴）表示数值。每个条形的长度对应着相应类别的数值大小。对于 Y Bar（垂直条形图），其横轴表示不同的类别，纵轴表示数值大小；而 X Bar（水平条形图）则是横轴表示数值，纵轴表示不同的类别。

适用情形

1. **比较不同类别之间的数值大小**：例如比较不同产品的销售额、不同地区的人口数量等。
2. **展示数据的分布情况**：可以通过对连续变量进行分桶（binning），然后用条形图展示每个桶内的数据数量或频率。
3. **显示时间序列数据**：将时间作为类别，对应的数值作为条形的长度，展示数据随时间的变化趋势。
4. **多变量对比**：可以在同一个条形图中展示多个变量，通过不同的颜色或样式区分，便于进行对比分析。

8.直方图是一种用于展示数据分布的统计图表，以下是其相关信息：

- **概念**：对于数据表中的某一列连续型数值变量，通过“分桶”（也叫“分箱”）的方式将其划分为若干个区间，然后统计每个区间内数据的频数或频率，并用矩形的高度来表示相应的频数或频率，从而直观地展示数据在各个区间的分布情况。例如，对于新股发行的市盈率这一连续型数值变量，可以用 bucket 函数进行“分桶”，生成一个新的离散变量，再以此绘制直方图。
- **适用情形**：主要用于展示连续型数据的分布状况，帮助人们了解数据在不同取值范围内的集中程度和变化趋势。比如在分析股票数据时，可以通过绘制市盈率的直方图，了解不同市盈率区间内股票的数量分布情况，进而把握市场整体的估值水平分布特征。
- **实现代码**：在使用 PerspectiveWidget 实现直方图时，假设数据表中有连续型

数值变量“pe”（市盈率），可以先使用 bucket 函数对其进行“分桶”操作，如 bucket("pe", 10) 表示将“pe”按照每 10 个单位进行分桶，生成新的离散变量（如命名为“bucket_pe”）。然后将“bucket_pe”设置为 Y Bar 视图的横轴（Group By），把任意其他一系列变量用 count（计数）函数汇总后设置为纵轴（Y Axis），这样就可以得到直方图。

对于数据表中的某一系列连续型数值变量（比如新股发行的市盈率 pe），我们经常希望观察其分布（distribution）。可以用 bucket 函数对连续变量进行“分桶”（比如表达式 bucket("pe", 10)），生成一个新的离散变量（比如命名为 bucket_pe），然后把离散变量设置为 Y Bar 的横轴（Group By），把任意其他一系列变量用 count（计数）函数汇总，设置为纵轴（Y Axis）。这样看到的就是直方图。“分桶”在有的地方也叫“分箱”（bin），其粒度大小需要根据数据适当调节。

8. Line Chart 的概念、适用情形以及实现代码

概念

折线图是一种常见的数据可视化图表类型，通过将数据点连接成折线来展示数据随时间或其他连续变量的变化趋势。它使用横轴表示时间或其他连续变量，纵轴表示数值，数据点之间通过线段连接，直观地呈现数据的变化情况。

适用情形

1. **时间序列分析**：用于展示数据在一段时间内的变化，例如股票价格走势、销售额随季度的变化等。
2. **趋势分析**：帮助观察数据的上升或下降趋势，以及趋势的变化情况，例如气温随月份的变化趋势。
3. **多组数据比较**：可以同时展示多组数据的变化趋势，便于比较不同组数据之间的差异和关系，例如不同产品的销量变化趋势对比。
4. **预测和趋势预测**：基于历史数据的折线图可以为预测未来数据提供参考，例如根据过去几年的销售额预测未来的销售额趋势。

9. 以下是 2020 至 2025 年期间影响 A 股或 IPO 的一些重大国内外财经事件：

2020 年：

新冠疫情全球爆发，对全球经济和金融市场造成巨大冲击。各国出台大规模经济刺激政策，以稳定经济增长。中国推出一系列支持企业发展的政策，包括减税降费、增加信贷投放等，为企业度过难关提供了支持，也为后续的 IPO 市场稳定奠定了基础。

新《证券法》正式实施，全面推行注册制，简化了企业上市流程，提高了 IPO 的

效率和透明度，为更多企业上市融资创造了条件。

2021 年：

全球经济逐步复苏，A 股市场表现活跃，IPO 数量和融资规模均有所增长。

北交所成立，聚焦服务创新型中小企业，为中小企业提供了更便捷的上市融资渠道，进一步完善了中国多层次资本市场体系。

2022 年：

俄乌冲突爆发，引发全球金融市场波动，大宗商品价格上涨，对全球经济和 A 股市场产生一定影响。

中国继续推进资本市场改革，加强对上市公司的监管，提高上市公司质量。同时，加大对科技创新企业的支持力度，推动相关企业在科创板等板块上市。

2023 年：

全球经济面临增长放缓压力，A 股市场也受到一定影响，但 IPO 市场仍保持一定活跃度。

中国政府出台一系列政策促进经济复苏，包括扩大内需、推动产业升级等，为企业发展提供了机遇，也支撑了 IPO 市场。

2024 年：

A 股 IPO 企业主要扎堆于工业、科技和材料行业，行业集中度显著升高。新“国九条”及资本市场“1+N”系列政策文件重磅出台，新股发行坚定秉持以质量为先、毫不动摇支持科技创新的宗旨，发行上市制度的包容性、适应性持续增强。

全球货币政策出现调整，对跨境资本流动和 A 股市场资金面产生一定影响。部分企业根据市场环境调整了 IPO 计划。

2025 年：

国内经济增速预计保持稳健，财政赤字率适度扩大，央行降息，为企业上市营造了良好的宏观环境。

监管层持续发力，整治市场乱象，优化革新 IPO 审核流程。同时，加大对新兴产业和高新技术企业的支持力度，降低企业上市门槛，提高审核效率。

这些事件对 A 股 IPO 市场的融资额和市盈率可能产生了多方面的影响。例如，政策的支持会促进企业上市，增加 IPO 融资额；市场的波动和不确定性可能会影响投资者情绪，进而影响市盈率等估值指标。通过结合这些事件分析示例数据中的融资额与市盈率变化情况，可以更好地理解数据背后的现实背景和原因。

10.Scatter Plot 的概念、适用情形以及实现代码

概念

散点图是一种用于展示两个数值型变量之间关系的可视化图表。在散点图中，每个数据点对应一对数值，分别表示在两个变量上的取值。通过观察散点的分布情况，可以判断两个变量之间是否存在某种相关性，例如正相关（散点呈现上升趋势）、负相关（散点呈现下降趋势）或无相关性（散点分布杂乱无章）。此外，还可以通过将更多变量映射为散点的颜色、大小、符号等属性，来展示更多维度的数据信息。

适用情形

相关性分析：用于研究两个数值型变量之间的关系，例如身高和体重之间的关系、销售额和广告投入之间的关系等。

数据分布展示：可以展示数据在二维空间中的分布情况，帮助了解数据的集中趋势、离散程度等特征。

多变量分析：通过将更多变量映射为散点的其他属性，可以在一个图表中展示多个变量之间的关系，例如在散点图中用颜色表示不同类别，用大小表示另一个数值变量等。

异常值检测：容易发现数据中的异常点，即与其他数据点明显不同的点，这些异常点可能代表特殊的情况或数据错误。

散点的分布如果特别不均匀，则意味着变量单位可能有问题，或者需要经过变换（比如取对数）

散点的分布如果杂乱无规律，则意味着 X 与 Y 没有相关性

散点的分布如果看起来能够拟合成一条直线（即回归线，regression），则意味着 X 与 Y 具有正的或负的相关性，意味着可能存在某些规律

散点图上可以进一步体现更多的变量维度，比如可以把更多变量映射为散点的不同颜色 (Color)、大小 (Size)、符号 (Symbol)、标签 (Label)、提示框 (Tooltip) 等