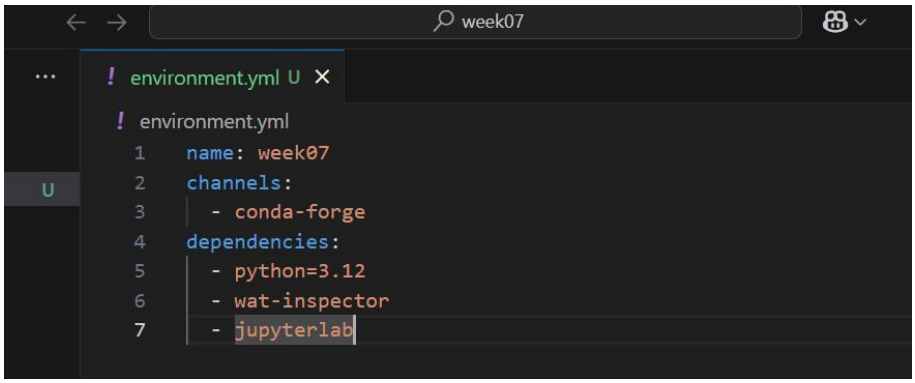
第七周作业笔记

一、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

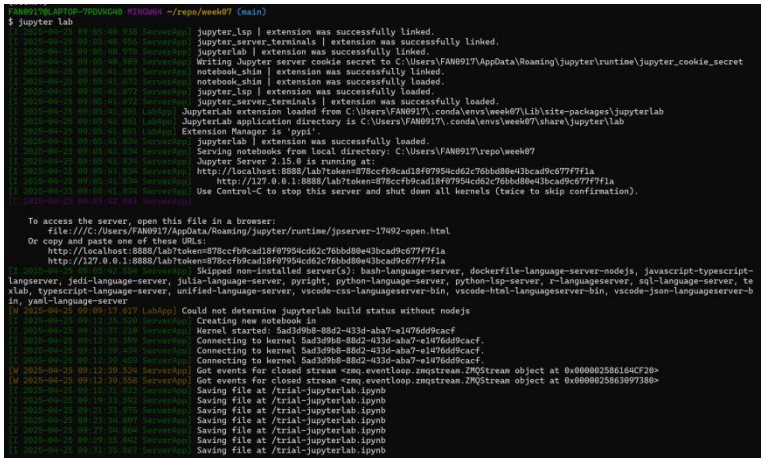
```
(base) FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo
$ git clone git@gitcode.com:jiemoduner/week07.git
Cloning into 'week07'...
remote: Enumerating objects: 5, done.
remote: Counting objects: 100% (5/5), done.
remote: Compressing objects: 100% (5/5), done.
remote: Total 5 (delta 0), reused 5 (delta 0), pack-reused 0 (from 0)
Receiving objects: 100% (5/5), 8.45 KiB | 1.21 MiB/s, done.
```

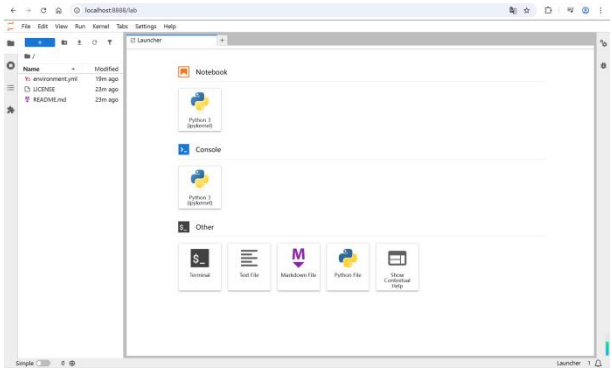
二、用 VS Code 打开项目目录，新建一个 environment.yml 文件，指定安装 Python 3.12 和 jupyterlab，然后运行 conda env create 命令创建 Conda 环境



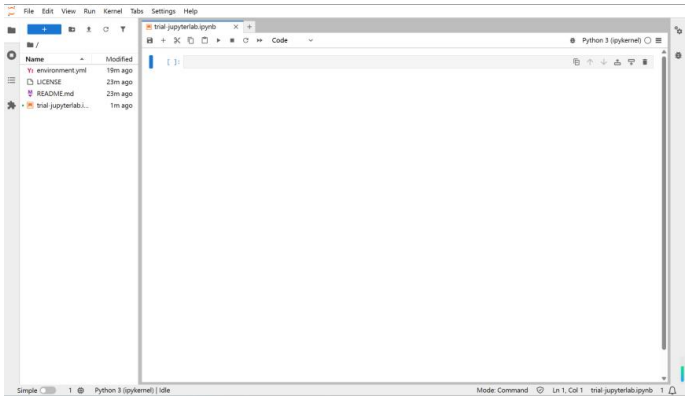

```
(base) FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo/week07 (main)
$ conda env create
D:\GQ\anaconda3\Lib\argparse.py:2006: FutureWarning: `remote_definition` is deprecated and wil
`conda env create --file=URL` instead.
  action(self, namespace, argument_values, option_string)
Retrieving notices: ...working... done
Channels:
 - conda-forge
 - https://repo.anaconda.com/pkgs/main
 - https://repo.anaconda.com/pkgs/r
 - https://repo.anaconda.com/pkgs/msys2
Platform: win-64
Collecting package metadata (repodata.json): done
Solving environment: done

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

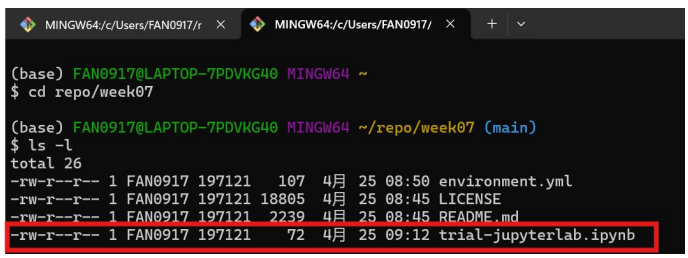
三、在项目目录下，运行 jupyter lab 命令，启动 后端（Backend）服务，在浏览器里粘贴地址访问 前端（Frontend）页面



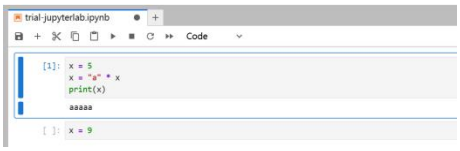


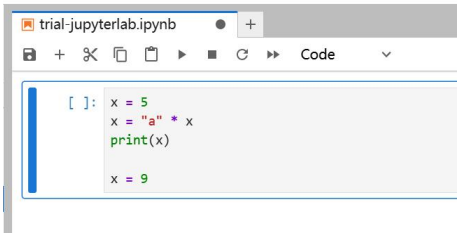
四、在 JupyterLab 页面里，新建一个 Notebook，改名为 trial-jupyterlab.ipynb





- 功能尝试——按 Shift+M 合并，按 z 撤销



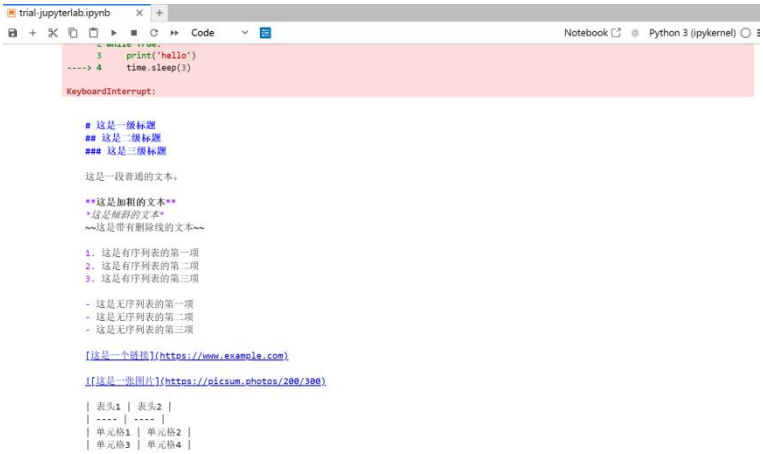


- 功能尝试——单元格最后一行如果是 表达式 (expression) 且运行后返回的对象不是 None，则计输出 (Out)，否则只计输入 (In)，序号为 i 的输出，可以用 _i 变量来引用

```
[2]: x = 5
     x = "a" * x
     print(x)
     x = 9

     aaaaa

[3]: x = 5
     x = "a" * x
     print(x)
     x = 9
     x * 3

     aaaaa
[3]: 27

[4]: a = [4,2,1]
     a

[4]: [4, 2, 1]

[5]: a.pop()

[5]: 1

[6]: a

[6]: [4, 2]

[7]: print(a)

     [4, 2]
```

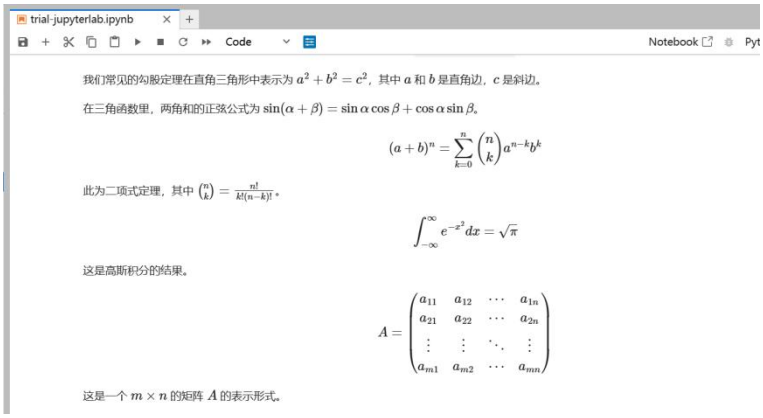
- 单元格 (Cell) 序号为 * 表示代码运行中，尚未返回，按 ii 可以打断 (KeyboardInterrupt) (类似于终端的 Ctrl+C)

- 用豆包生成一段示例 Markdown 代码，复制粘贴进 Markdown 单元格，运行以呈现 (Render)



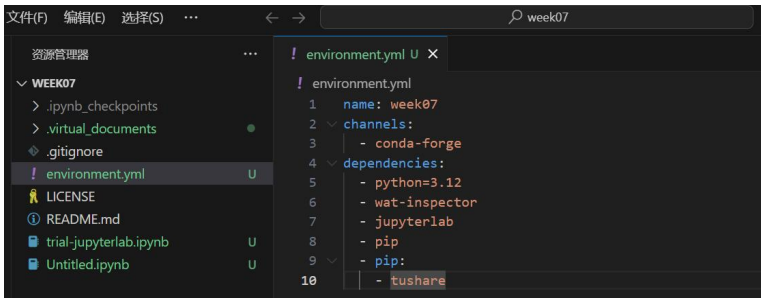
- 用豆包生成一段示例 HTML 代码，复制粘贴进 Markdown 单元格，运行以呈现 (Render)；注意不支持 CSS

- 用豆包生成一段示例 LaTeX 数学公式代码，复制粘贴进 Markdown 单元格，运行以呈现 (Render)；注意要用 $ (行内模式) 或 $$ (整行模式) 包围



五、通过 tushare 软件包下载保存一些数据

- 修改 environment.yml 文件，添加 pip: tushare (注意，conda-forge 没有收录 tushare，只能从 PyPI 安装，参考) 依赖项，运行 conda env update 更新 Conda 环境



➢ 在 IPython 提示符下，运行下面的 Python 代码设置 Tushare Token

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


(base) FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo/week07 (main)
$ python
Python 3.12.7 | packaged by Anaconda, Inc. | (main, Oct  4 2024, 13:17:27) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> import tushare as ts
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ModuleNotFoundError: No module named 'tushare'
>>> quit()

(base) FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo/week07 (main)
$ conda activate week07
(week07)
FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo/week07 (main)
$ python
Python 3.12.10 | packaged by conda-forge | (main, Apr 10 2025, 22:08:16) [MSC v.1943 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> import tushare as ts
```

➢ 按 Ctrl+D 结束前面的 IPython 进程，再重新启动一个新的 IPython 进程，运行下面的 Python 代码向 Tushare 服务器请求 IPO 新股列表 数据，并保存在本地

```
In [6]: pro.new_share()
Out[6]:
      ts_code sub_code  name  ipo_date issue_date   amount  market_amount  price     pe  limit_amount   funds  ballot
0   603014.SH   732014  威高血净  20250508       None   4114.0            0.0   0.00   0.00          1.10   0.000    0
.00
1   301595.SZ   301595  太力科技  20250508       None   2707.0            0.0   0.00   0.00          0.65   0.000    0
.00
2   688755.SH   787755  汉邦科技  20250507       None   2200.0            0.0   0.00   0.00          0.50   0.000    0
.00
3   301636.SZ   301636  泽润新能  20250428       None   1597.0          455.0  33.06  17.57          0.45   5.279    0
.00
4   920068.BJ   920068  天工股份  20250428       None   6000.0         4200.0   3.94  14.98        255.00   2.364    0
.00
...       ...      ...   ...       ...        ...      ...            ...   ...    ...           ...     ...  ...
1995  300777.SZ   300777  中简科技  20190506   20190516   4001.0         3601.0   6.06  22.98          1.10   2.425    0
.04
1996  603267.SH   732267  鸿远电子  20190430   20190515   4134.0         3721.0  20.24  16.50          1.60   8.367    0
.03
1997  600989.SH   730989  宝丰能源  20190430   20190516  73336.0        66002.0  11.12  22.07         22.00  81.550    0
.25
1998  300778.SZ   300778   新城市  20190425   20190510   2000.0         2000.0  27.33  22.99          2.00   5.466    0.
02
1999  002953.SZ   002953  日丰股份  20190424   20190509   4302.0         3872.0  10.52  16.34          1.70   4.526    0
.03

[2000 rows x 12 columns]
```

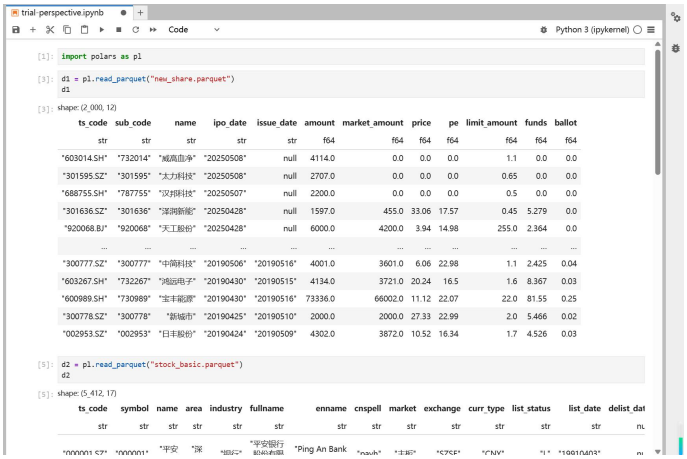
➢ new_share 接口只需要 120 积分，如果你有 2000 积分，可以采用与上面类似的方法访问 stock_basic 接口，并将数据保存为 stock_basic.parquet 文件 (注意，需要指定 fields 参数获取全部字段)。如果积分暂时不够，可以在终端运行下面的命令，从我们开源的 课程仓库 下载数据文件到你的本地

```
(base) FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo/week07 (main)
$ rm stock_basic1.parquet

(base) FAN0917@LAPTOP-7PDVKG40 MINGW64 ~/repo/week07 (main)
$ ls -lh
total 582K
-rw-r--r-- 1 FAN0917 197121  154  4月 28 10:24 environment.yml
-rw-r--r-- 1 FAN0917 197121  19K  4月 25 08:45 LICENSE
-rw-r--r-- 1 FAN0917 197121 118K  4月 28 10:28 new_share.parquet
-rw-r--r-- 1 FAN0917 197121 2.2K  4月 25 08:45 README.md
-rw-r--r-- 1 FAN0917 197121 423K  4月 28 10:48 stock_basic.parquet
-rw-r--r-- 1 FAN0917 197121 8.1K  4月 28 09:14 trial-jupyterlab.ipynb
-rw-r--r-- 1 FAN0917 197121   72  4月 28 08:56 Untitled.ipynb
```

六、通过 perspective-python 软件包查看 polars.DataFrame 数据，实践交互式可视化

➢ 调用 polars.read_parquet 函数，分别读取磁盘 (disk) 中的 new_share.parquet 文件和

stock_basic.parquet 文件，得到内存 (memory) 中的 polars.DataFrame 对象，命名为 d1 和 d2



➢ 数据可视化效果