

案例作业：股票研究 Agent —— 从代码到备忘录（24 小时）

限时 **24 小时**（预计有效投入 8–12 小时）。本作业不提供任何数据——**找数据本身就是题目的一部分**。我们看重思路与严谨性：未完成的功能不扣分（前提是诚实列明），但**核心链路必须能跑通**。

一、任务

搭建一个**股票研究 Agent**：输入任意一只股票代码，自动完成“取数 → 计算 → 分析 → 成文”全链路，输出一份**研究备忘录**。

```
python run.py --ticker <任意股票代码>
```

一条命令跑通全流程，是本作业的硬性验收标准。**验收时我们会用你没有见过的股票代码现场实测**（含市场上的“麻烦”标的），请不要针对特定股票写死任何逻辑。

市场任选其一：A 股或美股（二选一即可，需在设计文档中说明选择理由与数据源）。

二、Agent 必须覆盖的四个环节

■ 1. 取数（公开免费数据源）

- 自行选择数据源（如 akshare、yfinance 或其他公开渠道），获取该股票的行情、基本估值与财务数据；
- 数据范围自定，但备忘录中**每个数字都要标注数据截至日期（as-of date）与来源**；
- 数据源不可用、字段缺失是常态——你的处理方式是考察点，不是意外；
- 快照与离线重跑（硬性要求）**：Agent 运行时须把取到的原始数据自动落盘为快照（格式自定），并支持从快照**离线重跑**全流程（如 `--offline` 参数）。评分时我们会用你的快照复现你的数字（保证公平与可复现），实时取数模式用于现场实测；
- 如无可用的 LLM API，可在收题时申请由我们提供额度。

■ 2. 计算（确定性代码，非 LLM）

产出一组**可验证的量化事实**，至少覆盖：

维度	示例（不限于此）
价格与动量	区间收益率、相对基准超额、距 52 周高/低点位置
波动与流动性	年化波动率、日均成交额、极端交易日统计
估值位置	当前估值指标及其 历史分位数 （窗口与口径自定并说明）
财务质量	你认为最有信息量的 3—5 个指标

计算口径（复权方式、年化约定、分位数窗口、缺失值处理） **必须在输出中可见**。口径错误在本作业中的权重高于功能数量。

■ 3. 分析与成文（LLM）

- 用 LLM 将量化事实组织成一份**研究备忘录**（Markdown 或 PDF），建议结构：概览 → 量化事实 → 估值位置 → 风险信号 → 数据质量与局限 → 免责声明；
- **硬性约束：备忘录中的所有数字必须来自计算层的输出，LLM 不得自行生成或改写任何数字。** 你需要说明（最好用机制保证）这一点如何做到；
- 备忘录是"分析"，不是荐股：结论须与证据强度匹配，并附局限声明。

■ 4. 稳健性

- 对"麻烦"标的（长期停牌、上市不满一年、被特别处理、财务数据缺失等）：**允许输出降级**（如部分模块标注"数据不足"），**不允许崩溃或输出错误数字**；
- 数据接口失败要有基本的重试/报错策略；
- 用一页以内说明：你自己如何测试这个 Agent（测了哪些情形、怎么知道数字算对了）。

三、提交物（24 小时截止时）

#	提交物	要求
1	设计文档	≤ 2 页：架构与数据流图、LLM 与规则的分工、你的取舍
2	代码 + README	依赖、运行命令、数据源说明；克隆后 10 分钟内可跑起来
3	两份样例备忘录	用你自选的 2 只股票由 Agent 自动生成（非手写）
4	自测说明 + 未完成清单	各半页；诚实的未完成清单不扣分，被现场问出来的漏项扣分
5	AI 协作说明	≤ 1 页：附 1 个关键 prompt + 一个真实的"AI 出错—发现—修正—防再犯"案例（一个错都没发现会被视为验证不足）

答辩（提交后另约 30 分钟）：现场对我们指定的 1—2 只新股票实跑（预算 10 分钟），其余时间围绕设计取舍与计算口径问答。

四、说明

- LLM 环节使用你自己可用的任意模型（API 或本地模型均可）；若确实无法接入，可用规则模板替代成文并说明原设计，但集成完成度会相应影响评分；
- 允许并鼓励使用 AI 编程工具（Claude Code / Cursor 等）；答辩中你须能解释每一处专业口径与代码核心逻辑；
- 评分权重：**计算口径的严谨性 > 数据工程的稳健性 > 分析内容 > 功能数量**。一个只有 6 个指标但口径全对、边界稳健的 Agent，远好于 20 个指标但复权都没处理的 Agent；
- 本作业输出仅用于能力评估，不构成任何投资建议；请勿在备忘录中使用夸大或诱导性表述。