Data Science Project Schedule

9/26 22:00 Form the teams. Each team consists of at most 5 members and at least 3 members. The team leader turns in a one-page slide with team member names.

10/10 22:00 Turn in the project proposal in 6 pages of slides containing the project's possible title and abstract. The last page records the date, place, topics, and attendees of each group meeting. At least three group meetings should be held before 10/11.

10/11 The team leader presents the project proposal within 8 minutes. The presentation language could be English or Chinese.

10/24 22:00 Each team member submits a paper (search from google scholar, after 2017) that is closely related to your project problem.

11/07 22:00 Turn in the second version of your project proposal.

11/08 One of the team members (not the team leader) presents the project proposal within 8 minutes. The presentation language could be English or Chinese.

Two required items of the second project report.

1. formal literature review.

Provide at least 4 closely related recent papers (after 2017) from Google Scholar. Each paper should be reviewed by one team member. Two sides for each paper review state (a) the purpose of the research, (b)the main methods and approaches, (c) the relation to your project, (d) the key accomplishment of the paper

1. Date summarization and visualization. Provide many summary statistics to describe the data. Make histogram, bar chart, correlation plot, pair plot, box plot, time series plot,…(any useful plots) to visualize the data.

The other detailed items of the second report will be provided later.

Will update the schedule soon.

Grading principles

This is a project for the data science course, and the goals of the expected teaching effect are as follows

**Graduate Program Learning Goals** (goals covered by this course are indicated by checks):

|  |  |  |
| --- | --- | --- |
| ✓ | 1 | Graduate students should be able to appreciate data analysis approaches and to present research findings/ results effectively in speaking and in writing. |
| ✓ | 2 | Graduate students should be able to integrate different functional areas in solving data analysis problems. |
| Graduate students should be able to analyze data effectively and to recommend effective statistical methods. |
|  | 3 | Graduate students should be able to demonstrate leadership skills as a data analysis team leader. |
| Graduate students should be able to identify ethical dilemmas and to determine necessary courses of action. |
|  | 4 | Graduate students should possess a global data science perspective and an awareness of the global business. |
| ✓ | 5 | Graduate students should be able to coordinate actions and solve problems jointly with other members of a professional team. |

1. The content of the report slide is clearly described.
2. Prepare your report on stage and explain it clearly.
3. Give full play to the benefits of team division of labor and cooperation to solve problems together.
4. The topic content and presentation should conform to Jeff Hammerbacher’s Model of Data Science Work.
5. There must be sufficient review and comparison of relevant field knowledge (literature).
6. Proposal competitions focus on creativity, value, feasibility, and data verification.

The data available from FinTech Space

◼ 申請使用之API項目：(供應商名稱\_提供的API/服務名稱)

(請參考玖、數位沙盒實證項目清單填寫)

(A1)供應商:臺灣證券交易所。API/服務項目:上市即時行情。

(A2)供應商:臺灣期貨交易所。API/服務項目期貨/選擇權即時行情。

(A3)供應商:時報資訊。API/服務項目股票歷史資料和三年基金歷史資料。

(A4)供應商:台灣經濟新報。API/服務項目高風險股票篩選模組和實價登錄模組。

(A5)供應商:全曜財經資訊。API/服務項目CMoney Taiwan ESG Rating TOP 100。

(A6)供應商:意藍資訊。API/服務項目OpView社群口碑。

◼ 創新實證應用範圍：

智能理財: 彈性整合上述申請項目API (A1)~(A5)的資料內容，以機器學習模型和深度學習演算法進行大數據分析，建構股票、期貨、不動產投資之高階智能理財(Robo-Advisor or Smart financial management) Advance Integrated-API模組(AI-API)。

市場價值評估及信貸評估: 在申請項目(A6)設定特定目標族群，例如人氣100名網紅、人氣100名電商、追蹤其社群資料的每周變化，評估其社群媒體市場價值和持續影響力，計算信用價值評等(credit value score)和市場價值評等(market value score)，建立social networking evaluation API (SoNet-API)。

Data Science Project Instruction 2023/10/04

#==================================================================

The project proposals that are not based on the data from FinTech Space will be approved when the study goal is consistent with the **data science work** of the **Hammerbacher’s model.**

The progress of the project should follow the time schedule of this course.

#=================================================================

To identify your target problem, you should first gain **domain knowledge on the topic and on the data source.**

Step 1: Search web pages, ask ChatGPT, read papers, and discuss with teammates.

Step 2: Try to instrument (select and organize) data sources provided by FinTech Space.

Step 3: Repeat Step 1 and Step 2

Step 4: Write down the project proposal

Step 5: Repeat Step 1 – Step 3, modify the project proposal.

#=================================================================

In your proposal, the following items should be included.

The first page gives the course name, instructor name, and team member names.

List the team leader.

The Second page: Project title and target problems

What target problems do your team want to solve?

The third page: Importance and Novelty

Why the target problem is important? To whom it is valuable?

What parts are new?

The fourth page: Data description  
 What kinds of data will you use to solve your target problems?

The fifth page: Reference

List all the references you used. (web pages, ChatGPT, papers, thesis, …)

The sixth page: Group meeting records

The last page records the date, place, topics, and attendees of each group meeting. At least three group meetings should be held before 10/11.