

巨集程式與資料分析應用

W5. 生成式AI 輔助程式開發

黃家耀老師

Department of Transportation & Logistics Management
National Yang Ming Chiao Tung University



單元1：巨集程式及應用

- 第一週：Excel 功能與VBA基礎
- 第二週：Excel VBA進階
- 第三週：儀表板與資料視覺化
- 第四週：生成式AI (ChatGPT) 輔助程式開發
- 第五週：Term Project 1

學習目標

1. 瞭解 ChatGPT 的應用及限制
2. 學會運用 ChatGPT 輔助程式開發

學習資源

- [GPT best practices](#)
- [Prompt 的技術：ChatGPT 提問的四大層次 L1 ~ L2 \(上集\)](#)
- [Prompt 的技術：ChatGPT 提問的四大層次 L3 ~ L4 \(下集\)](#)
- [The way of Prompting](#)
- [ChatGPT 圖文混合提問 9 種超能力：程式、教育、學習到行銷應用](#)
- [ChatGPT 幫我不懂程式碼也能寫出番茄鐘待辦清單工具，實測心得](#)
- <https://www.youtube.com/@HungyiLeeNTU>

生成式AI

什麼是生成式AI？

- 一種人工智能模型，能夠生成新的數據或內容，可用應用於自然語言生成、圖像生成、音樂創作等。
- 其原理是基於大量訓練數據，學習數據分佈，然後生成與訓練數據相似的新數據。
- 基於大型語言模型（ Large Language Models，LLM ），可包含數十億個參數的的人工神經網絡。



<https://www.youtube.com/watch?v=MTjb1ps8iVs>

ChatGPT 可以用來做什麼？

- 聊天
- 翻譯
- 摘要、萃取、總結、重寫
- 寫故事、講笑話、規劃旅遊行程、提供菜單及食譜....
- 數據分析、畫圖、寫程式
- ~~寫作業、寫論文~~
- ...

使用生成式AI的注意事項

- 心態：擁抱科技、與AI共存
- 思維：結合AI工具及其他資源，解決複雜問題
- 預期：理解AI的能力和限制，設定合理的預期
- 責任：不完全依賴，為自己的作品及產出負責

國立陽明交通大學因應生成式AI之指引及教學建議（1120524）

ChatGPT 申請帳號

<https://chat.openai.com/> 申請免費（或付費）帳號

- GPT-3.5，訓練資料截至2021年9月

付費帳號（ChatGPT Plus）功能

- GPT-4，訓練資料截至2022年1月
- Browse with Bing（網絡瀏覽）
- Advanced Data Analysis
- Plugins
- DALL-E-3

I ONCE HEARD THAT THERE IS A PROGRAMMER



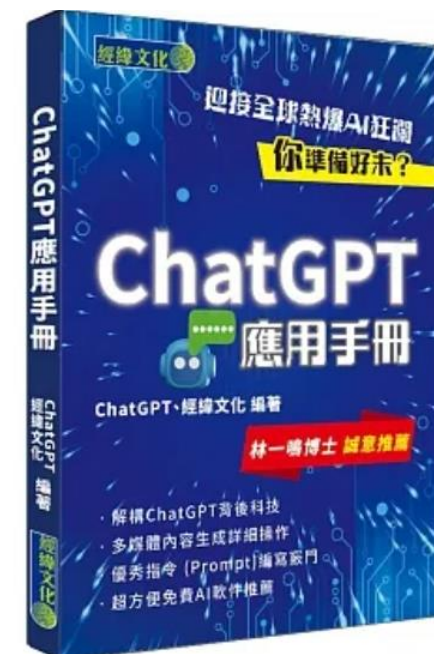
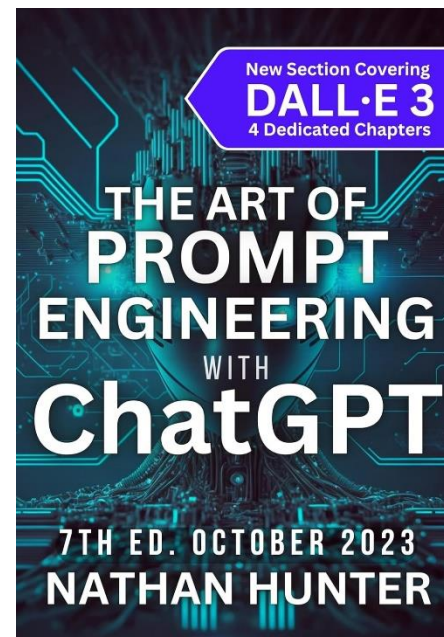
THAT WRITES CODE WITHOUT CHATGPT



使用技巧

Prompt Engineering (咒語)

- 分析需求：輸入、輸出、具體需求



使用技巧：角色扮演

- 指定模型扮演特定角色，可獲得特殊的回應或特定風格回應
- 例如指定法律、醫學專家給出專業意見

Initiation
啟始設定

設定角色，專業背景與知識

你現在是作家，你的專業是...

Instruction
工作指令

主要請ChatGPT要完成的工作

回覆郵件、總結、回答問題...

Input data
參考資料

與上下文相關的資料

郵件內容、文案參考、情境描述...

Specifiers
限定詞

格式、語氣、限制

以CSV格式撰寫，用專業的語氣，在500字以內...

You are a banking professional.

Answer the following E-Mail:

[EMAIL CONTENT]

Format the E-Mail in Markdown

啟始設定
設定角色

工作指令
回覆郵件

參考資料
郵件內容

限定詞
Markdown

使用技巧：提供例子

- Zero-Shot：不提供任何上下文或先前例子，直接向模型提問。
- One-Shot：提供一個相關例子或語境，以引導模型的回答。
- Multi-Shot：提供多個例子，以更精確地引導模型。

使用技巧：提供例子

- 提供任務說明，讓ChatGPT 學習規則，提升成功率

Language Models are Few-Shot Learners

<https://arxiv.org/pdf/2005.14165.pdf>

The three settings we explore for in-context learning

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 cheese => ..... ← prompt
```

One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← example
3 cheese => ..... ← prompt
```

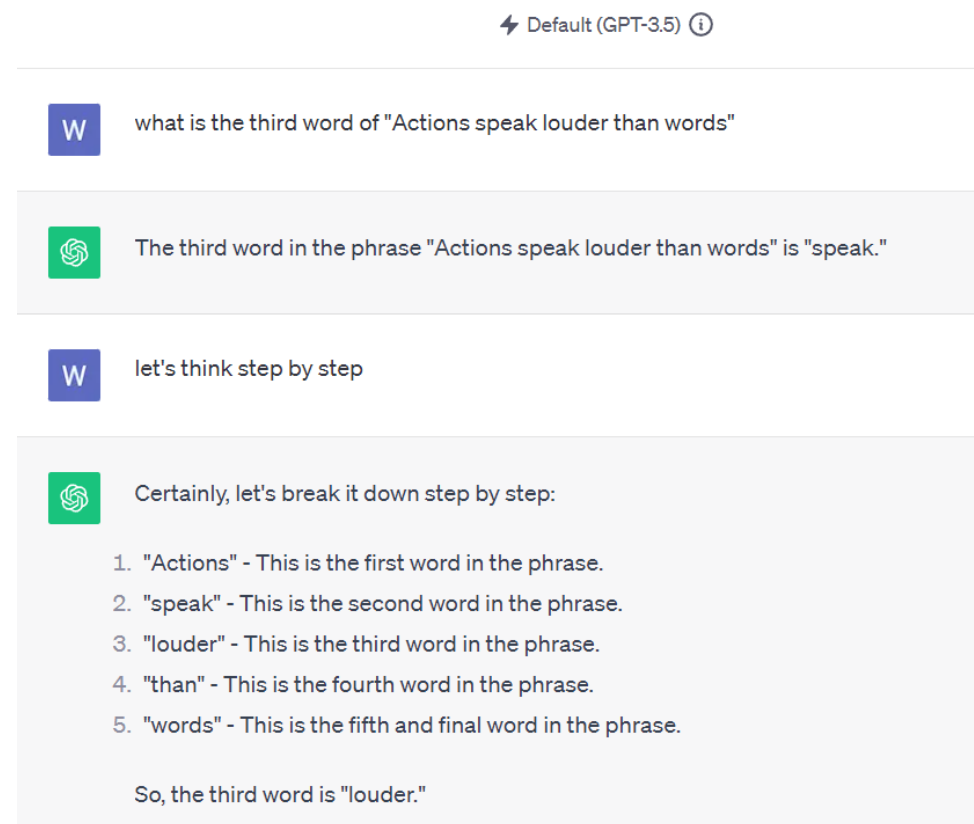
Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese => ..... ← prompt
```

使用技巧：提供引導

- ChatGPT 是基於 LLM，在語言理解及翻譯有巨大優勢，較不擅長邏輯及運算推理。
- Zero-shot 加入「**Let's think step by step**」會大幅提高答對的機率。
- 如涉及多步驟的推理問題，需要進一步提供引導性的問題，協助問題拆解。



練習1：行程規劃

- 請設計新竹遊的一天行程
- 5分鐘
- 嘗試提供一些限制條件，例如交通工具、景點喜好（歷史文化、美食、夜店...）
- 行程中是否有不合理的地方？

練習2：文獻格式修正

- 論文類作業需要引用文獻，但不同資料來源的格式都可能不相同，需要進行轉換：

原格式

- Greenberg, H., "An Analysis of Traffic Flow", Operations Research, Vol.7, pp. 78-85, Appendix A.21, 1959.

要求格式

- Greenberg, H., (1959年)「An Analysis of Traffic Flow」, Operations Research, Vol.7,第78~85頁, Appendix A.21.

練習2：文獻格式修正

請將以下文獻轉換至要求格式

- Greenberg, H., "An Analysis of Traffic Flow", Operations Research, Vol.7, pp. 78-85, Appendix A.21, 1959.
- Tian, Z. Z., Balke, K., Engelbrecht, R., and Rilett, L., "Integrated Control Strategies for Surface Street and Freeway Systems", Transportation Research Record, Vol. 1811, pp. 92-99, 2002.
- Jia, B., Jiang, R., and Wu, Q. S., "Traffic Behavior near an Off Ramp in the Cellular Automaton Traffic Model", Physical Review E, Vol. 69, No. 056105, pp. 130-135, 2004.
- Hagen, L., Lin, P.-S., and Fabregas, A., "A Toolbox for Reducing Queues at Freeway Off-Ramps", Center for Urban Transportation Research, Florida Department of Transportation, University of South Florida, pp. 44-59, Report BD 544.87-A-10, 2006.

練習2：解答

W

我有一些參考文獻，請您幫忙由格式1轉換為 格式2。

=====

格式1

Greenberg, H., "An Analysis of Traffic Flow", Operations Research, Vol.7, pp. 78-85, Appendix A.21, 1959.

格式2

Greenberg, H., (1959年) 「An Analysis of Traffic Flow」, Operations Research, Vol.7,第78~85頁, Appendix A.21.

參考文獻

Greenberg, H., "An Analysis of Traffic Flow", Operations Research, Vol.7, pp. 78-85, Appendix A.21, 1959.

Tian, Z. Z., Balke, K., Engelbrecht, R., and Rilett, L., "Integrated Control Strategies for Surface Street and Freeway Systems", Transportation Research Record, Vol. 1811, pp. 92-99, 2002.

Jia, B., Jiang, R., and Wu, Q. S., "Traffic Behavior near an Off Ramp in the Cellular Automaton Traffic Model", Physical Review E, Vol. 69, No. 056105, pp. 130-135, 2004.

Hagen, L., Lin, P.-S., and Fabregas, A., "A Toolbox for Reducing Queues at Freeway Off-Ramps", Center for Urban Transportation Research, Florida Department of Transportation, University of South Florida, pp. 44-59, Report BD 544.87-A-10, 2006.

<https://chat.openai.com/share/73788131-382d-49e4-82ac-a3e927f00268>

練習3：資料分析

請運用作業#1 的資料，分析進出口貨物的年度銷售金額變化。

- 因有字數輸入限制，先運用前100筆資料進行分析。

- 步驟

1. 資料整理：將前100筆原始資料貼到ChatGPT，並指示需要整理為可分析的格式
2. 資料計算：計算進口貨物、出口貨物的年度銷售金額
3. 輸出格式：將分析結果整理為表格

練習3：步驟1 資料整理

先試試看 ChatGPT 的能力，是否能理解問題題意

⚡ Default (GPT-3.5) ⓘ



某一經銷電腦週邊產品之跨國公司，銷售產品有主機板、處理器、記憶體、磁碟機、顯示卡、音效卡、網路卡等七種產品，並於銷售文件中紀錄每一筆交易的日期等資訊。某日因公司登錄系統故障，導致資料格式錯亂。

以下為100筆資料，將資料整理為可分析狀態
資料格式將整理為：年、月、產品來源、產品名稱、銷售額

====

4.5萬元 國產 9月 2001年 網路卡
進口 7月 29.8萬元 磁碟機 2005年
2011年 國產 12月 1.6萬元 音效卡
國產 2月 2016年 音效卡 41萬元
5月 78.5萬元 國產 處理器 2008年
2010年 音效卡 91.4萬元 進口 7月
2006年 記憶體 3月 國產 36.5萬元
2016年 國產 音效卡 50.6萬元 11月

練習3：步驟1 資料整理

格式有符合我們的需求，有看懂問題

< 3/3 >




以下是整理後的資料，包括年、月、產品來源、產品名稱和銷售額：

- 2001年 9月 國產 網路卡 4.5萬元
- 2005年 7月 進口 磁碟機 29.8萬元
- 2011年 12月 國產 音效卡 1.6萬元
- 2016年 2月 國產 音效卡 41萬元
- 2008年 5月 國產 處理器 78.5萬元
- 2010年 7月 進口 音效卡 91.4萬元
- 2006年 3月 國產 記憶體 36.5萬元
- 2016年 11月 國產 音效卡 50.6萬元
- 2002年 3月 國產 主機板 33.2萬元
- 2008年 2月 國產 音效卡 45.4萬元
- 2009年 9月 國產 磁碟機 46.2萬元
- 2010年 5月 進口 處理器 87.3萬元
- 2010年 6月 國產 處理器 86.5萬元

練習3：步驟2 資料計算

提出需求，如無想法也可以請ChatGPT 提供 (brainstorming)

W 請協助提供基本統計分析

 以下是根據您提供的資料進行的基本統計分析：


總銷售額：

- 總銷售額：約1,896.7萬元

月份銷售統計：

- 1月：總銷售額約255.7萬元
- 2月：總銷售額約224.1萬元
- 3月：總銷售額約161.2萬元
- 4月：總銷售額約204.3萬元

W 我想知道每年進口及出口的銷售金額變化，請協助分析

 以下是每年進口和出口銷售金額的變化分析：

進口銷售金額：

- 2001年：0元
- 2002年：51萬元
- 2003年：7.3萬元
- 2004年：14.5萬元
- 2005年：73.9萬元
- 2006年：94.5萬元
- 2007年：27.8萬元

練習3：步驟3 輸出格式

將資料整理為表格，最好明確指示各欄位的內容



請提供表格，欄位為：年份、進口銷售金額、出口銷售金額

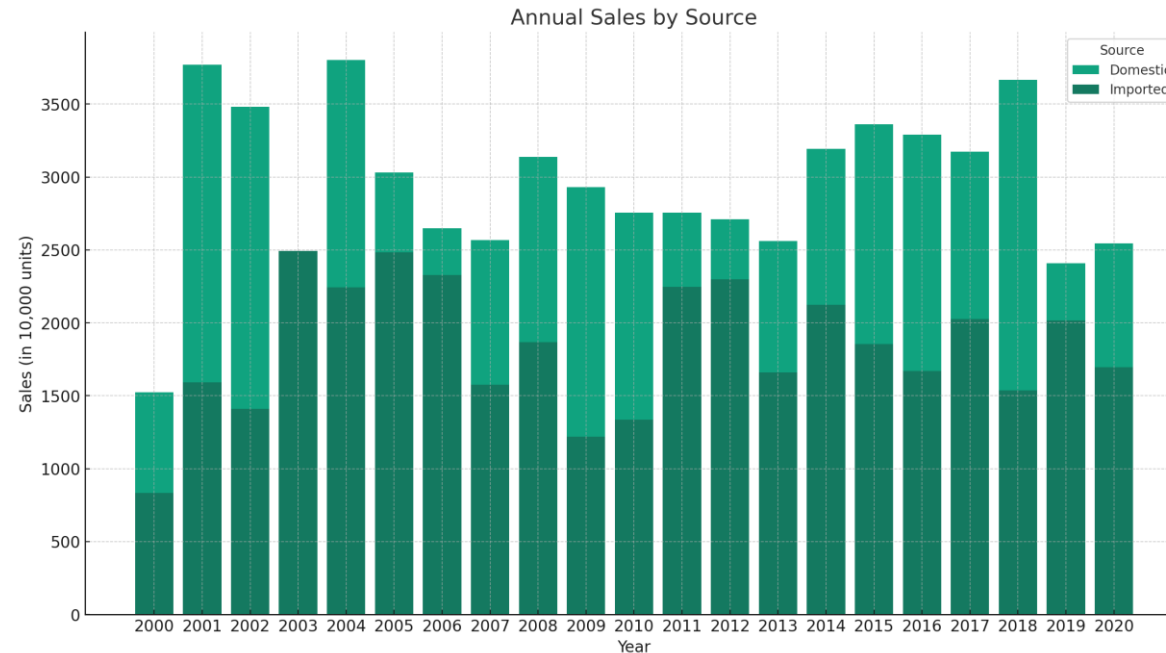


以下是每年的進口銷售金額和出口銷售金額的表格：

年份	進口銷售金額 (萬元)	出口銷售金額 (萬元)
2001年	0	227.5
2002年	51	84.2
2003年	7.3	177.1
2004年	14.5	177.9
2005年	73.9	296.2
2006年	94.5	291.8
2007年	27.8	315.6

練習3：步驟3 輸出格式 (Optional)

ChatGPT Plus 的 Advanced Data Analysis 可以畫圖並提供 Python 程式碼



練習4：協助寫程式

嘗試運用ChatGPT 完成作業#1 的巨集程式撰寫

某一經銷電腦週邊產品之跨國公司，銷售產品有主機板、處理器、記憶體、磁碟機、顯示卡、音效卡、網路卡等七種產品，並於銷售文件中紀錄每一筆交易的日期等資訊。某日因公司登錄系統故障，導致資料格式錯亂。

要如何輸入？

	A	B	C	D	E
1	4.5萬元	國產	9月	2001年	網路卡
2	進口	7月	29.8萬元	磁碟機	2005年
3	2011年	國產	12月	1.6萬元	音效卡
4	國產	2月	2016年	音效卡	41萬元
5	5月	78.5萬元	國產	處理器	2008年
6	2010年	音效卡	91.4萬元	進口	7月
7	2006年	記憶體	3月	國產	36.5萬元
8	2016年	國產	音效卡	50.6萬元	11月

練習4：直接輸入問題

- 直接將題目問題輸入到ChatGPT，不加修改

⚡ Default (GPT-3.5) ⓘ



某一經銷電腦週邊產品之跨國公司，銷售產品有主機板、處理器、記憶體、磁碟機、顯示卡、音效卡、網路卡等七種產品，並於銷售文件中紀錄每一筆交易的日期等資訊。某日因公司登錄系統故障，導致資料格式錯亂。

1. 請撰寫一巨集程序Sub DataCleaning()，將資料整理為可分析狀態

- ✖ 資料格式將整理為：年、月、產品來源、產品名稱、銷售額，另為方便後續計算，將數值之單位移除，並加入表頭；
- ✖ 於相同檔案內建立一個新的工作表並儲存結果。

前8行資料如下

====

4.5萬元 國產 9月 2001年 網路卡
進口 7月 29.8萬元 磁碟機 2005年
2011年 國產 12月 1.6萬元 音效卡
國產 2月 2016年 音效卡 41萬元
5月 78.5萬元 國產 處理器 2008年
2010年 音效卡 91.4萬元 進口 7月
2006年 記憶體 3月 國產 36.5萬元
2016年 國產 音效卡 50.6萬元 11月

練習4：直接輸入問題

- 會得到一個有模有樣的 VBA 程式碼
- 執行有些小bug
 - 工作表：“原始資料” -> “工作表1”
- 無法完全解決問題



以下是用VBA巨集程序撰寫的Sub DataCleaning()，可以將資料整理成可分析的狀態，並建立一個新的工作表儲存結果：

```
vba Copy code

Sub DataCleaning()
    Dim SourceSheet As Worksheet
    Dim DestinationSheet As Worksheet
    Dim LastRow As Long
    Dim i As Long

    ' 假設資料在名為"原始資料"的工作表中
    Set SourceSheet = ThisWorkbook.Sheets("原始資料")

    ' 在同一個工作簿中建立一個新的工作表
    Set DestinationSheet = ThisWorkbook.Sheets.Add
    DestinationSheet.Name = "整理後資料"

    ' 複製表頭到新工作表
    SourceSheet.Rows(1).Copy Destination:=DestinationSheet.Rows(1)

    ' 找到最後一行資料的行數
    LastRow = SourceSheet.Cells(SourceSheet.Rows.Count, "A").End(xlUp).Row

    ' 遍歷每一行資料進行整理
    For i = 2 To LastRow
        Dim OriginalData As String
        Dim DataParts() As String
        Dim CleanedData As String

        ' 讀取原始資料
        OriginalData = SourceSheet.Cells(i, 1).Value
```

<https://chat.openai.com/share/6cc0e5cf-e9f5-4d8c-9ccf-71cfe04e4004>

練習4：提供明確指示

- 加入『咒語』：**Let's think step by step.**
- 提供更明確的說明及指示：

資料儲存在“工作表1”，每一筆資料共包含5個數據，
第一筆資料在 Excel 的 A1、A2、A3、A4、A5；
第二筆資料在 Excel 的 B1、B2、B3、B4、B5，如此類推
共有2000筆資料

練習4：拆解問題，引導求解方式

- 拆解問題，引導求解步驟

需要將每一筆資料的5個數據，判斷是屬於“年”、“月”、“產品來源”、“產品名稱”、“銷售額”的哪一個欄位

年的最後一個字是“年”，將“年”移除

月的最後一個字是“月”，將“月”移除

銷售額的最後是“萬元”，將“萬元”移除

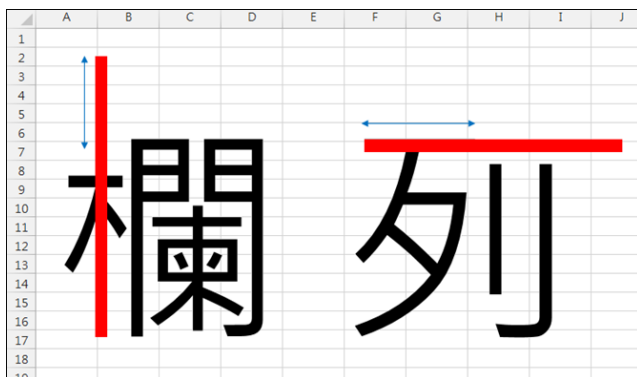
產品來源會是“國產”或“進口”

產品名稱會是主機板、處理器、記憶體、磁碟機、顯示卡、音效卡、網路卡 7項其中之一

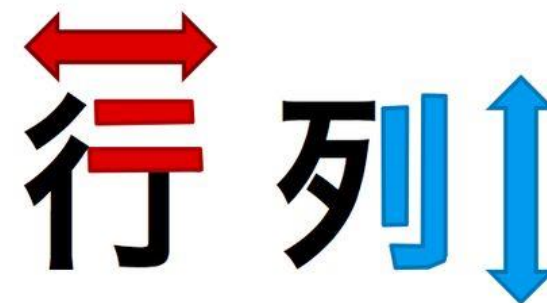
<https://chat.openai.com/share/59756e41-572f-4ca3-ba8d-0ade1995868c>

練習4：中文用詞及亂碼

- 如果遇到亂碼，加入咒語：代碼中如果有中文，請用 #zh-tw
- 注意中文用詞差異
 - 台灣：直行橫列 = 行 (column) 列 (row)
 - 中國：直列橫行 = 行 (row) 列 (column)



Source: <https://pic.pimg.tw/isvincent/1422060927-872711917.png>



<https://i.imgur.com/EOh0vuS.jpg>

課後調查 W5

- 請同學離開前，協助填寫課後調查問卷。謝謝你的參與！
- 此調查將有助於我們更好地規劃及調整課程內容。
- 連結在 Padlet