

# 2023教育大數據專題製作

## 初賽提案構想書

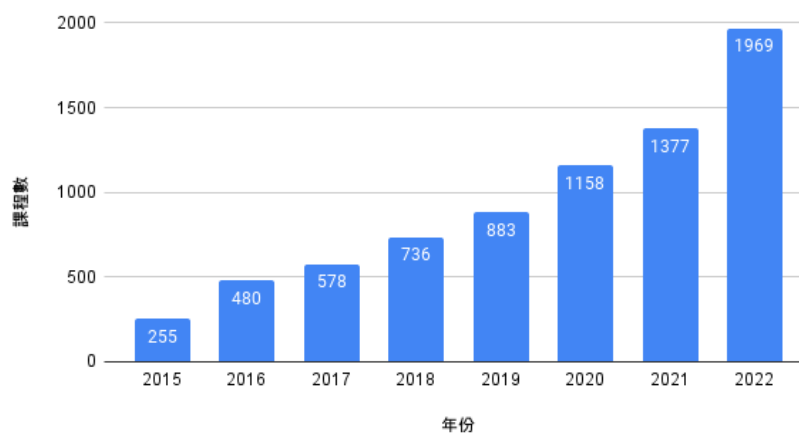
團隊名稱：愉快的小夥伴們

議題：暑假學習真的好嗎-線上教學的成效評估與改善建議

### 一、提案議題

在這個資訊發達的世代，線上教育成為了一種新型的教育方式，此外在疫情的影響之下，數位化課程因此蓬勃發展。Statista的資料表示，從2015年至2022年間，FutureLearn 線上提供的活躍獨特課程數量從255堂上升至1969堂，翻了將近7.7倍。此外SELLCOURSES ONLINE的報導指出，全球線上學習市場規模從3150億美元增長到了1兆美元，許多網路名人甚至是補教名師也在疫情期間，先後製作出多元的線上課程提供使用者觀看與學習，但線上課程的學習效率是有效的嗎？根據政大機構典藏中的線上學習成效影響因素模式之探討，線上教學如果互動性不夠高，或是教材呈現方式不佳，有可能導致線上學習的成效嚴重下滑。綜上所述，我們想藉由這次的報告來談論、研究及分析，找出提升學生使用線上課程學習成效的方法。

課程數量表



(課程數量增長圖，資料來源：

<https://www.statista.com/statistics/1094391/number-of-futurelearn-courses/>)

影響要素	說明
資訊要素-資訊環境穩定性	電腦環境與網路連線設施是否符合學習者需求。
資訊要素-網站使用便利性	使用者是否可以熟悉介面設計與功能架構。
教學要素-課程內容	學習內容是否具備其意義，且是否可以引起學習者興趣，與教材設計是否能夠考慮遠端學習者需要。
教學要素-教材呈現方式	教材是否多元可以吸引學習者注意，以及學習者設備使用是否可以負荷教材。
教學要素-互動性	是否可以利用資訊科技所設計的互動模式，來提高學習動機與成效。
環境要素	學習者環境差異也影響學習之成效。
個人要素-資訊先備經驗	學習者的電腦及網路使用經驗與能力
個人要素-學習者本身特性	學習者對於課程使用是否具備主動、積極、內控之特性。

(線上教育的影響要素，資料來源：

政大機構典藏-線上學習成效影響因素模式之探討

[https://nccur.lib.nccu.edu.tw/bitstream/140.119/90355/1/79\(%20p1-21%20\).pdf](https://nccur.lib.nccu.edu.tw/bitstream/140.119/90355/1/79(%20p1-21%20).pdf))

## 二、動機與目的

身為疫情後的第一波受害者，我們曾被迫接觸線上教學。然而，當時線上教學品質參差不齊，不僅學生不適應線上上課，學習成效也不佳。即便老師師資優越，也可能因不熟悉設備而影響教學品質，使好的教育資源被浪費。這些因素導致我們的線上教學體驗不盡理想。

舉我們親身經驗來說，我們發現許很多同學容易在線上上課分心。可能是因為他們不習慣在無人監督的環境上課，且相較實體課程，線上教學的外在誘因更多，也就更容易讓人精神漫遊，因此不適合長時間的課程。這讓我們意識到，線上教學不能盲目搬遷實體教學方式，而需要適當調整，才能發揮其效益。

在實體課程中，許多同學會透過筆記提升專注力和知識記憶，甚至於去借別人的筆記來看。然而，在線上課程中，學生未必會記得做筆記。有些同學在觀看影片時可能忘記停下來做筆記，或是因為分心錯過

重要的知識點，這讓我們感到可惜。

幾年過去，線上教育發展迅速，進步神速。藉由更符合人性的軟體和線上平台，教學方提供更完整的配套措施，不僅提供一般的線上教學，還有即時性的線上測驗輔助課程，更有效地幫助學生吸收課程內容並檢測學習成效。我們對這令人振奮的進展感到激動，也讓我們燃起為改善線上教學環境貢獻一份心力的渴望。

透過線上教學平台的大數據資料，我們希望透過檢視學生的學習模式和成效看出一些端倪。同時，藉由親身經歷和觀察到的問題，分析學生在線上學習時可能遇到的困難和習慣。我們計劃透過分析學生的影片操作，找出最適宜的影片長度，提供給平台一些建議與數據佐證。此外，我們還會分析做筆記對學生的影響，比較有無筆記的學生在學習成效上的差異，以提供改善線上教學的建議。我們期望能實現讓學生愛上線上教學環境的願景，並使教育方達成其預期之學習成效的目標，同時宣洩我們對改善線上教育的渴望。

線上教學應該是突破距離和空間限制的創新教學方式。它不僅讓學生和教師無需為了學習而花時間通勤，還能夠在任何讓學生和教師感到舒適的地方進行學習，進而提升學習成效。甚至，透過科技輔助，讓教學變得更簡單而有效。因此，我們希望透過此競賽，藉由大數據分析線上教學的成效，找到改進方案，改善學生在線上空間中的學習模式，使線上教學成為學生自主學習的得力助手。

### 三、數據準備與處理

#### (一)、各月份瀏覽次數、成績收集與初步比對：

我們先大致預覽了資料內容，按照使用者在平台學習的月份(month)分成六月的在學期間組和七八月的放假期間自主學習組，分別比較觀看次數和作答情形，其中觀看次數由影片瀏覽序號(review\_sn)的出現次數多寡決定，直接的代表了線上教學的使用程度以及需求，同時，我們將使用者在平台中瀏覽影片時檢核點作答的結果(exam\_binary\_res)視為學習成效的指標。

由於兩者數據量不一致，所以我們取兩組的作答成績的平均值，並計算標準差和變異數，以提供更多參考的依據，透過這兩項數據，我們對比了線上教學在學期間和暑假中何時使用需求量較高，以及不同時間中使用者的學習成效比較。

#### (二)、線上教學操作行為整理與初步處理：

為了給予線上教學的修正建議，我們分析每個使用者於學習平臺瀏覽影片時影片操作的行為資料，紀錄每個同樣的影片(video\_item\_sn)不同使用者執行影片操作的行為名稱

(record\_plus\_view\_action)，以及影片操作的時間戳記(record\_plus\_view\_time)，並觀察影片中暫停(paused)，降低播放速度(slowdown)、提升播放速度(speedup)等按鍵是否多次被使用，以及多個使用者是否在影片中相近的時間點做出相同操作行為，來探討此一影片時長的適切性，需不需要在特定時間點分割影片，使影片長度符合使用者的專注力。

另外結合作答結果(exam\_binary\_res)和影片能力指標(indicator)，討論影片的內容是否過於艱深，導致學習成效不佳，並討論是否應該將該影片內容切分成不同的小主題，以方便使用者吸收。

### (三)、測驗結果與觀看次數收集與學習成效之函數生成：

最後，為了更進一步將最大化線上教學的成效，我們考慮使用者的行為和線上教學成效的關係，首先是作筆記，我們將影片操作的行為資料紀錄中的停下做筆記(note)作為變因，透過兩個組別的檢核點作答結果(exam\_binary\_res)，以及練習題作答的正確率(prac\_score\_rate)進行以下比較：

1. 同一使用者觀看同一影片時有做筆記前後檢核點作答結果(exam\_binary\_res)以及練習題作答的正確率(prac\_score\_rate)的差異。

2. 不同使用者中，有做筆記的檢核點作答結果(exam\_binary\_res)以及練習題作答的正確率(prac\_score\_rate)平均值是否高於所有使用者的平均值。這有助於研究做筆記之行為對於學習的成效的影響，並討論是否建議使用者在線上教學時搭配做筆記。

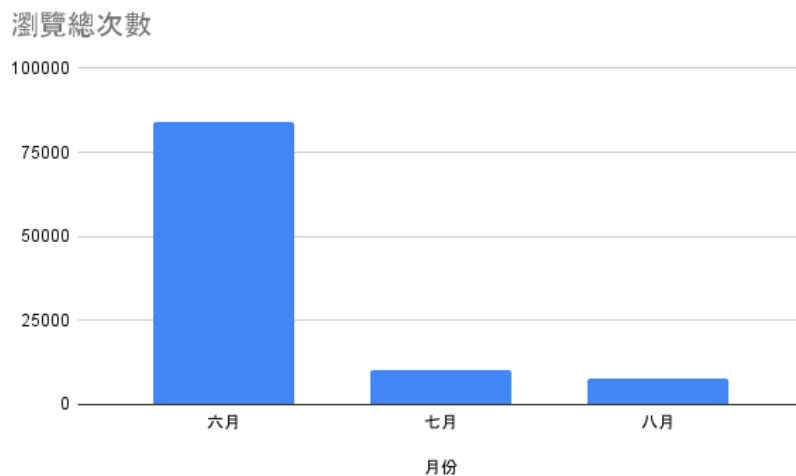
其次觀看次數，統計不同的使用者資料(PseudoID)中同一影片ID(video\_item\_sn)的出現次數，代表該使用者觀看此一影片的次數，對照觀看次數和檢核點作答結果(exam\_binary\_res)以及練習題作答的正確率(prac\_score\_rate)，將其結果按照資料量規格化，探討觀看次數和學習成效的函數關係，此流程同樣應用到不同能力指標(indicator)、資源的知識架構分類(classification)，以整理出不同種類型的影片在觀看多少次後會達到相對有效的學習成效，並提供各能力指標間預期學習成效的建議觀看次數。

變數類別	變數名稱	變數使用說明
使用者基本資料	month: 使用者使用平臺學習的月份	用以將資料分成六月的在學期間組以及七、八月的放假期間自主學習組
學習平臺 dp001 上的影片基本資料	indicator: 能力指標	判斷影片內容學習難易度的依據，視資料結果給出修改建議
	video_item_sn: 影片 ID	用以計算該影片的閱覽次數
使用者於學習平臺 dp001 瀏覽影片時的基本資料	review_sn: 影片瀏覽序號	用以計算使用者觀看影片的次數
使用者於學習平臺 dp001 瀏覽影片時影片操作的行為資料紀錄	record_plus_view_time: 執行影片操作的時間戳記	搭配影片操作的行為名稱，統計觀看過該影片的使用者行為，給出影片的修改建議
	record_plus_view_action: 影片操作的行為名稱	內容含有: paused-暫停、slowdown-降低播放速度、speedup-提升播放速度、replay-重看一次，搭配執行影片操作的時間戳記，統計觀看過該影片的使用者行為，給出影片的修改建議
使用者於學習平臺 dp001 瀏覽影片時進行檢核點作答時的資料紀錄	exam_video_exam_sn: 檢核點測驗記錄 ID	將練習題以及檢核點作答測驗紀錄和結果互相搭配，反映使用者對該影片的學習成效
	exam_binary_res: 作答結果	
使用者於學習平臺 dp001 上進行練習題作答的資料紀錄	prac_score_rate: 練習正確率	
	prac_binary_res: 每一題的作答結果 (1 答對, 0 答錯)	

#### 四、分析方法與流程

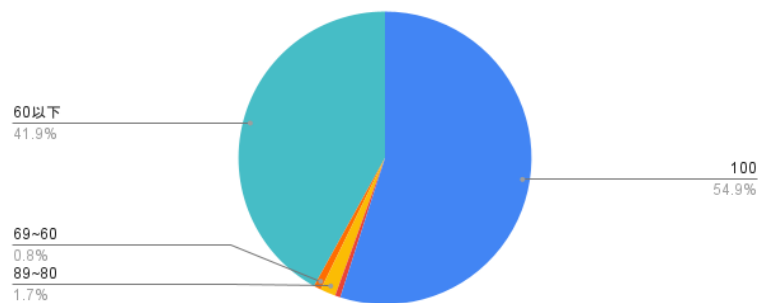
##### (一)、各月線上教學的需求以及使用教學分析：

藉由每個影片在各個月份中的瀏覽次數去觀察學生在不同月份的使用程度以及需求，此外從作答的人數以及答對率(prac\_score\_rate)與六月的數據進行比對，以此來分析學生在六月上課階段以及七、八月的放假階段的學習狀況以及線上教學的成效。



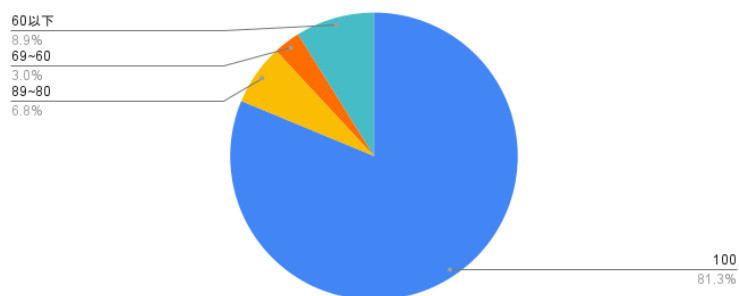
(瀏覽總次數圖)

六月各成績分布

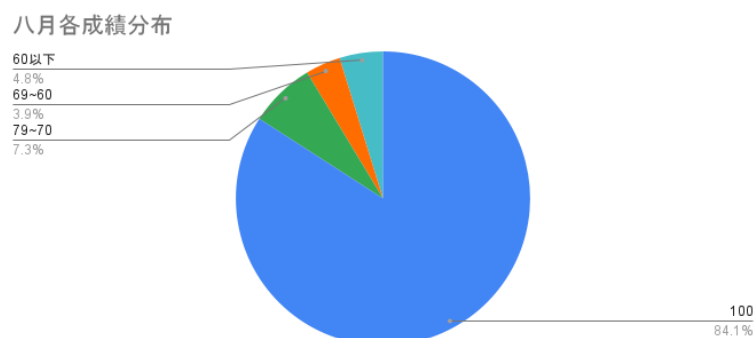


(六月各成績分布)

七月各成績分布



(七月各成績分布)

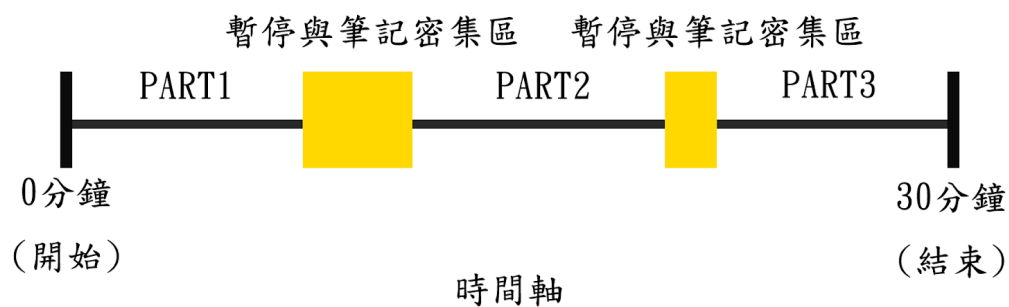


(八月各成績分布)

月份	平均	高標	底標
6	72.30	54.22	18.07
7	83.95	62.96	20.98
8	93.09	69.81	23.83

(二)、藉由學生按下暫停與作筆記之時間點做影片分布或是暫停建議：

藉由學生在觀看影片時的操作(record\_plus\_view\_action: 影片操作的行為名稱)，來觀察暫停與作筆記的時間密集度來分析在影片中的哪些時間段是學生經常按下暫停或是撰寫筆記的操作行為，以此來給予暫停或是提醒此時間段為重點之建議或是將這部影片在進行影片長度處理。



(影片暫停與筆記密集區圖)

(三)、藉由答題狀態來建議學生觀看影片的次數：

藉由每部影片學生的答題狀態(prac\_during\_time: 練習花費時間、prac\_score\_rate: 練習正確率)與觀看次數(review\_sn: 影片瀏覽序號)來分析學生的學習狀況、答題狀態，給予需要再觀看幾次的建議，或是在首次觀看此影片時給予建議觀看次數，使學生能先了解其知識點之難度。

利用所獲取之數據來分析權重與運算式，來對應練習正確率與觀看次數之關聯性，並在學生有提高正確率或是降低正確率時給予建議影片觀看次數。

權重的算法是將各個影片的prac\_during\_time: 練習花費時間、prac\_score\_rate: 練習正確率與review\_sn: 影片瀏覽序號做收集並且形成線性方程式與設定權重，使再觀看次數可以進行預測。

A因素 \* 權重1 + B因素 \* 權重2 + C因素 \* 權重3 = 需要再觀看次數

A因素: prac\_during\_time(練習花費時間)

B因素: prac\_score\_rate(練習正確率)

C因素: 觀看次數

(建議觀看次數運算式)

## 五、初步結果與預期效益

### (一)、在學期間使用率較高

時間上，將資料分為兩類，上學期間（6月）以及假期（7月、8月），根據數據資料結果顯示學生在在學期間的使用次數明顯比假期多上許多，以成績分布來看，假期高分的(大於59分的)成績分布比上學期間高出許多，基於這些資料我們初步推測可能的原因。

#### 1. 教師進行推動:

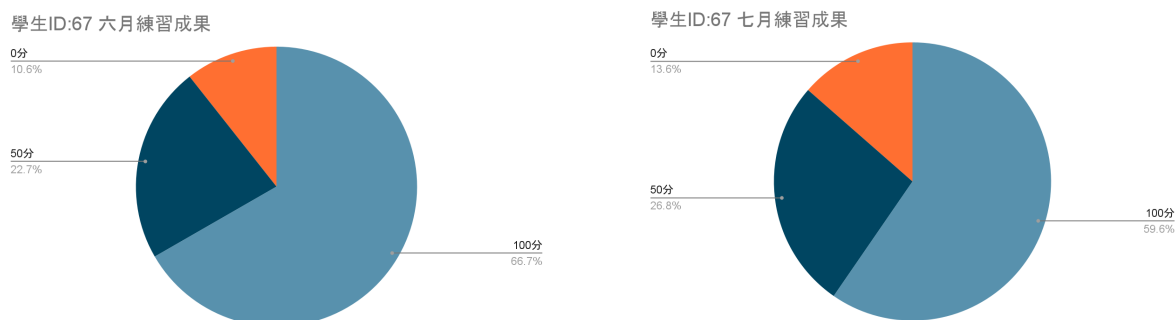
在上學期，由於教師的推動，學生們願意花時間參與線上課程。然而，進入假期時，由於缺乏教師的監督，線上學習成為考驗學生自主學習意願的鏡子。

#### 2. 學生的特性是否積極且主動:

透過對假期成績分布的分析方法，我們發現一種可能性：那些在學期間認真學習的學生，在假期仍然持續努力地進行題目練習，進而保持高分水平，使得高分族群維持穩定，而中、低分族群的百分比則有所下降。與此同時，我們也觀察到有一部分學生僅在假期進行學習，且其暑假成績相對較學期平均成績更為優秀。

以下圖為例，展示了同時在上學期和假期進行線上教學的學生所取得的練習成果。





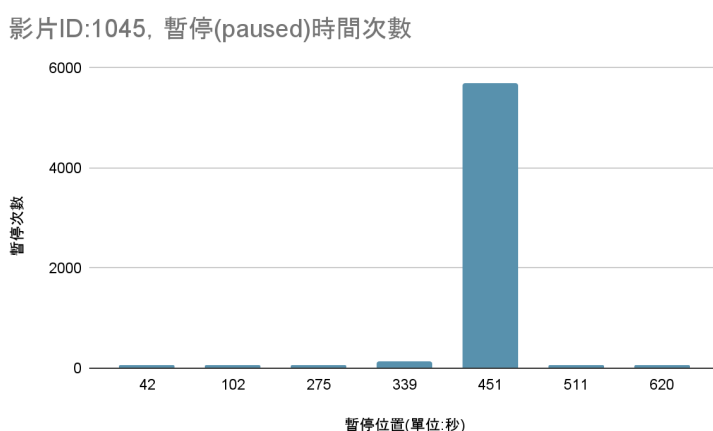
(學生練習成果比較圖)

從圖可以看出學生在學期間的練習成果相較不好並與整體數據有不合之情況，我們分析可能有以下原因。

1. 學生相較不情願花費更多時間和精力在線上學習上。
2. 學生本身未具有相對先備知識。
3. 學生可能對於其領域相較不主動、積極。

## (二)、按下暫停的時間點非常集中

根據資料資料分析的結果可以看出該部影片在哪個區段暫停次數最多次，也可以藉此查出停下來做筆記的時間，下圖以影片ID:1045為例。



(影片暫停位置圖)

透過觀察影片播放進度，我們可以發現在第451秒時有大量的暫停需求。由此可推斷，在影片的451秒處，內容可能相對較為困難，或者學生在該時段已經感到疲勞。

## (三)、改善的方向和建議

初步成效分析後，改善的方向可分成三類：

1. 藉由影片行為密集點給予建議：

首先，將影片從分析出的密集暫停點切成小段落，可提昇學生專注力及提供抄筆記的時間，可加入題目練習加強記憶點。

除了以單元來分影片的段落之外，我們考慮到了學生的專注度。研究指出，兒童的注意力時間會隨年齡增長而增加，3~5歲為5分鐘，5~7歲則是10分鐘，7~10歲增為20分鐘，而12歲以上到達30分鐘。在不同的年齡段有不同的專注時間，未來可以再用專注度將長度做更精細的分析，看哪種長度的影片下，學生答題成功率較高，作為此年齡下最合適的教材長度。

## 2. 筆記共享與作筆記建議：

記筆記是一種融會貫通的方法，對成績有顯著的影響，要求學生在某些較難的科目單元上傳筆記，可以讓他們透過手寫，將本章的重點劃記下來。不論是紙本還是數位版，在複習或忘記如何作答時，有個筆記與重點讓他們可以看，使學生不用再重新去把整部影片看完。

## 3. 影片再觀看次數建議：

某些單元本身較難理解，我們有發現很多人會將影片反覆拉回觀看某一片段，我們將那些高峰片段的答題率，用答題正確率、答題時間來設定一個次數。比如說，如果本章你只有60分，時間花了約30分鐘才做完，那我們就設定5次為你的影片理想觀看次數。

## 六、參考文獻：

SELLCOURSES ONLINE-100+ Incredible Online Learning Statistics in 2023

<https://sellcoursesonline.com/online-learning-statistics>

Statista-Number of active unique courses available online at FutureLearn from 2015 to 2022

<https://www.statista.com/statistics/1094391/number-of-futurelearn-courses/>

政大機構典藏-線上學習成效影響因素模式之探討

[https://nccur.lib.nccu.edu.tw/bitstream/140.119/90355/1/79\(%20p1-21%20\).pdf](https://nccur.lib.nccu.edu.tw/bitstream/140.119/90355/1/79(%20p1-21%20).pdf)