

國小生於線上學習平台的 練習分數和影片學習成效 之關聯與影響因素

第五組 江允菡 楊子萱

OUTLINE

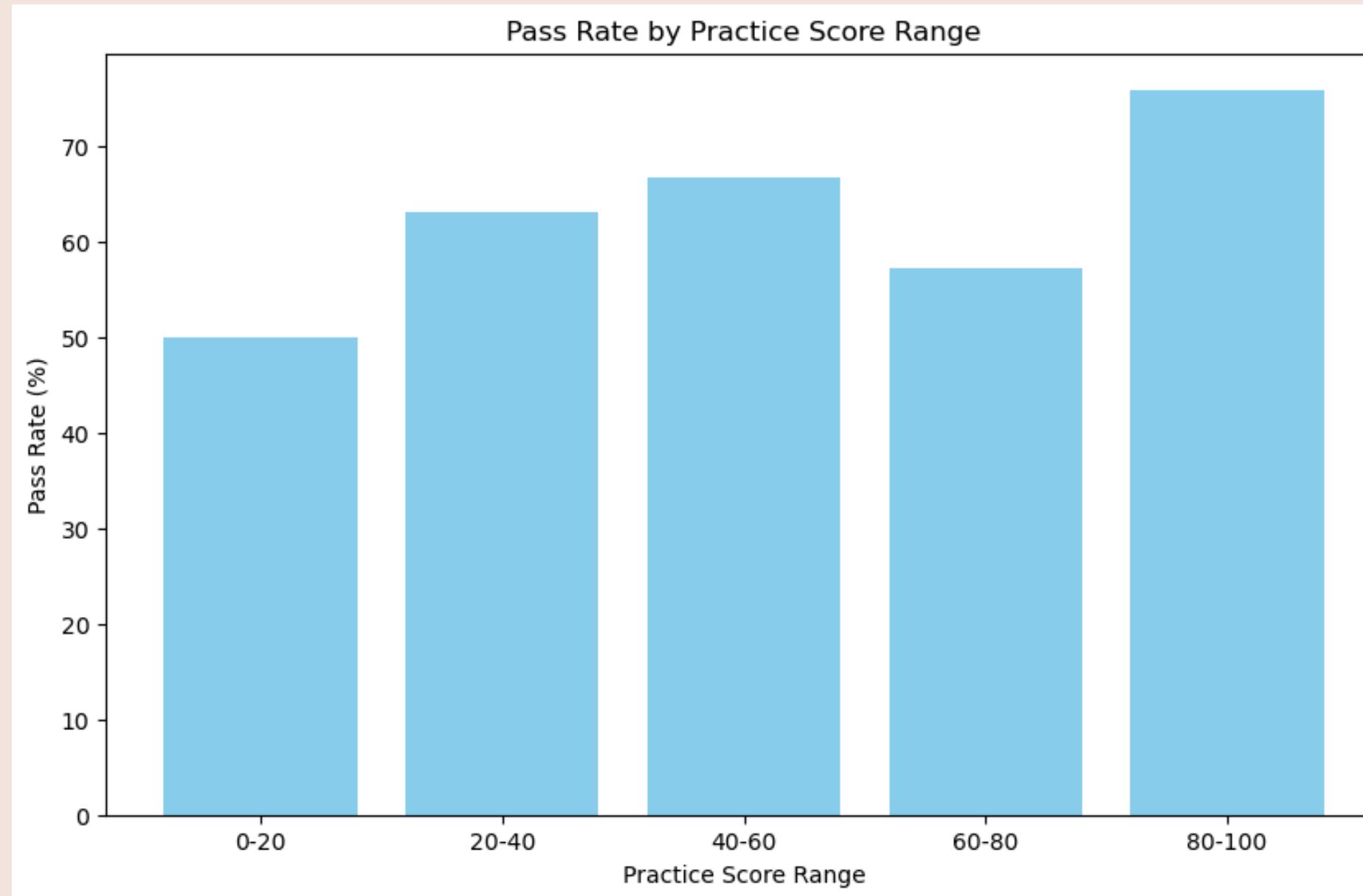


- INTRODUCTION
- METHOD
- RESULT
- CONCLUSION
- FUTURE WORK

INTRODUCTION



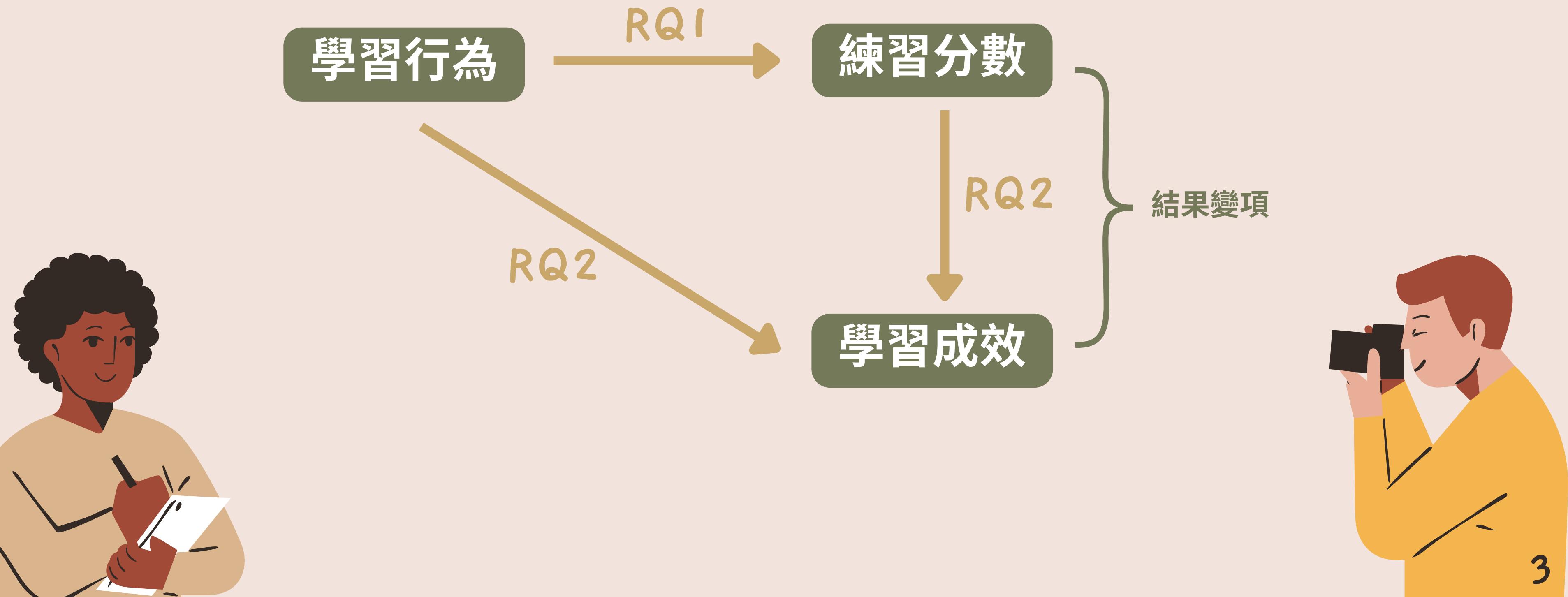
研究動機



測驗效果 (*testing effect*) 指出，反覆練習測驗能強化記憶與理解，提升學習表現。
(Akbulut, 2024)

然而由數據可見，練習分數與影片學習成效並非穩定正相關，這使我們思考是哪些因素影響了練習分數，使得它無法直接代表影片學習成效？

研究框架





研究問題

針對 DPOOL 平台

RQ1

影片觀看行為、練習時長
和學習時間點是否會影響
練習分數？

RQ2

學生的練習分數是否與影
片學習成效相關？

METHOD



資料集

針對 DP001 平台

RangeIndex: 353240 entries, 0 to 353239

Data columns (total 25 columns):

#	Column	Non-Null Count	Dtype
0	PseudoID	353240	non-null int64
1	dp001_indicator	272207	non-null object
2	dp001_video_item_sn	352666	non-null float64
3	dp001_question_sn	222956	non-null float64
4	dp001_review_sn	352666	non-null float64
5	dp001_review_start_timestamp	352666	non-null float64
6	dp001_review_start_time	352666	non-null object
7	dp001_review_end_timestamp	353240	non-null float64
8	dp001_review_end_time	352666	non-null object
9	dp001_review_total_time	352666	non-null float64
10	dp001_review_finish_rate	352666	non-null float64
11	dp001_exam_video_exam_sn	222956	non-null float64
12	dp001_exam_question_timestamp	170477	non-null float64
13	dp001_exam_ans_time	222956	non-null float64
14	dp001_exam_binary_res	222956	non-null float64
15	dp001_record_plus_view_time	333987	non-null object
16	dp001_record_plus_view_action	333987	non-null object
17	dp001_record_plus_timestamp	333987	non-null float64
18	dp001_item_plus_question_timestamp	222956	non-null float64
19	dp001_prac_date	229174	non-null object
...			
24	dp001_prac_binary_res	229174	non-null object



取所需欄位、移除空值、重複資料後，
共有 2387 筆資料

PseudoID	dp001_review_total_time	dp001_review_finish_rate	dp001_prac_during_time	dp001_prac_date	dp001_review_start_time	dp001_prac_score_rate	dp001_exam_binary_res
5	573.46	100	32	2022/9/28 21:20	2022-09-28 21:17:04	100	1
22	573.46	100	32	2022/9/28 21:20	2022-09-28 21:17:04	100	0
22	449.61	100	129	2022/9/15 21:16	2022-09-15 20:40:28	80	1
26	449.61	80	154	2022/9/15 21:23	2022-09-15 20:40:28	100	1
26	100.96	22	129	2022/9/15 21:16	2022-09-15 16:55:33	67	1

合併 data1, data2 後共 353240 筆資料

研究問題 1

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

因素：

a. 影片觀看行為

`dp001_review_total_time` (總瀏覽時間)

`dp001_review_finish_rate` (影片完成率)

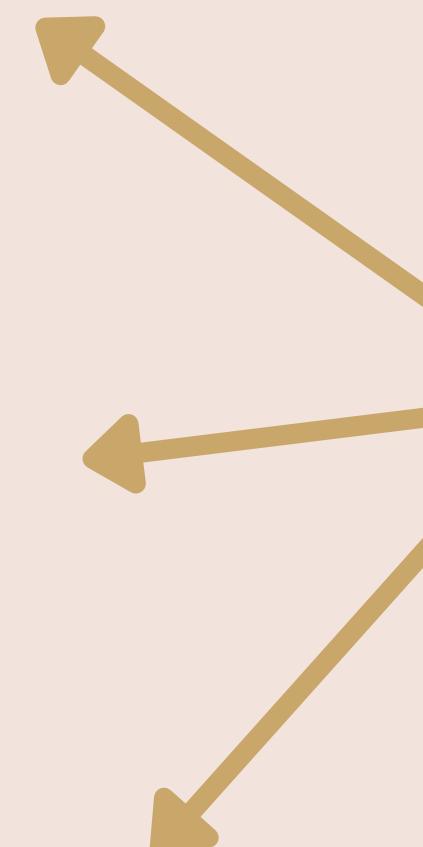
b. 練習時長

`dp001_prac_during_time` (練習時長)

c. 學習時間點

`dp001_prac_date` (練習時間)

`dp001_review_start_time` (觀看時間)



RangeIndex: 2387 entries, 0 to 2386			
Data columns (total 10 columns):			
#	Column	Non-Null Count	Dtype
0	PseudoID	2387	non-null int64
1	<code>dp001_review_total_time</code>	2387	non-null float64
2	<code>dp001_review_finish_rate</code>	2387	non-null float64
3	<code>dp001_prac_during_time</code>	2387	non-null float64
4	<code>dp001_prac_date</code>	2387	non-null object
5	<code>dp001_review_start_time</code>	2387	non-null object
6	<code>dp001_prac_score_rate</code>	2387	non-null float64
7	<code>dp001_prac_date_cat</code>	2387	non-null category
8	<code>dp001_review_start_time_cat</code>	2387	non-null category
9	<code>dp001_prac_score_rate_cat</code>	2387	non-null category

研究問題 1

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

觀看與練習的時間點：

- 將dp001_prac_date 與 dp001_review_start_time 轉換成早上 (06:00~11:59)、下午 (12:00~16:59)、晚上 (17:00~23:59) 與凌晨 (00:00~05:59)

	dp001_review_start_time_cat
morning	959
afternoon	861
evening	567

	dp001_prac_date_cat
morning	944
afternoon	912
evening	527
early morning	4

- 由於凌晨資料數過少，選擇將其資料點刪除

練習分數：

- 將dp001_prac_score_rate 轉換成高 (70~100)、中 (40~69) 和低 (0~39)

	dp001_prac_score_rate_cat
high	1304
medium	677
low	406

研究問題 2

練習分數與影片學習成效之關聯

主要分析欄位：

- dp001_prac_score_rate (練習分數)
- dp001_exam_binary_res (影片學習成效)

學習行為因子：

- dp001_review_total_time(總瀏覽時間)
- dp001_review_finish_rate(影片完成率)
- dp001_prac_during_time(練習時長)

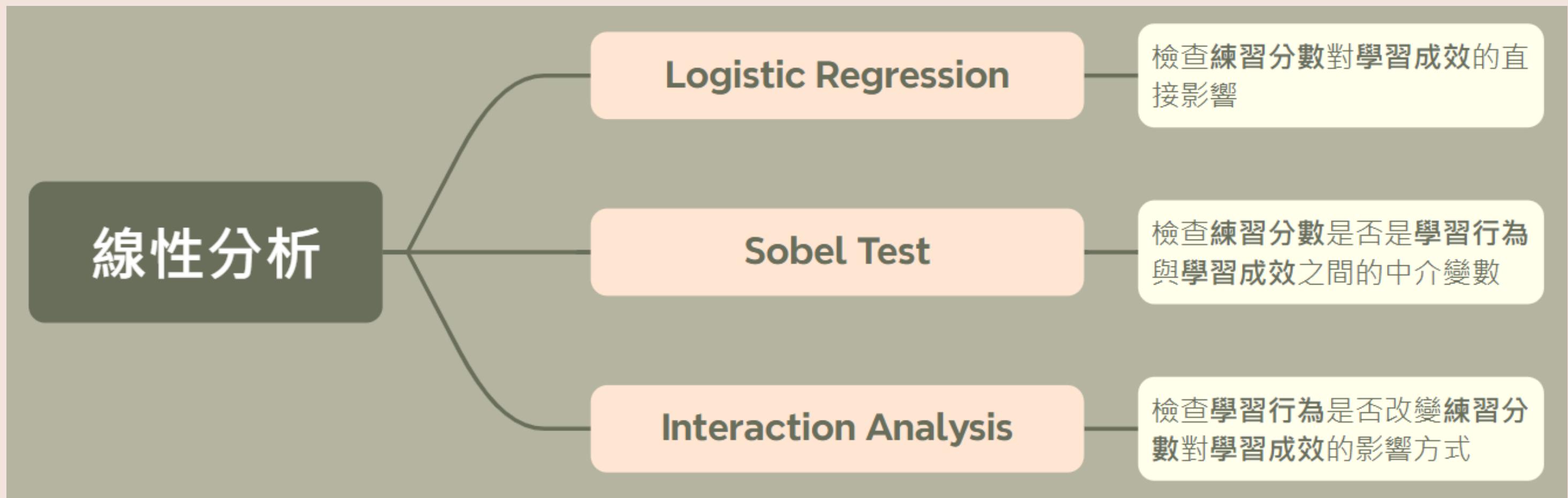


2個主要問題：

- 練習分數與影片學習成效是否相關？
- 哪些是影響影片學習成效的關鍵因素？

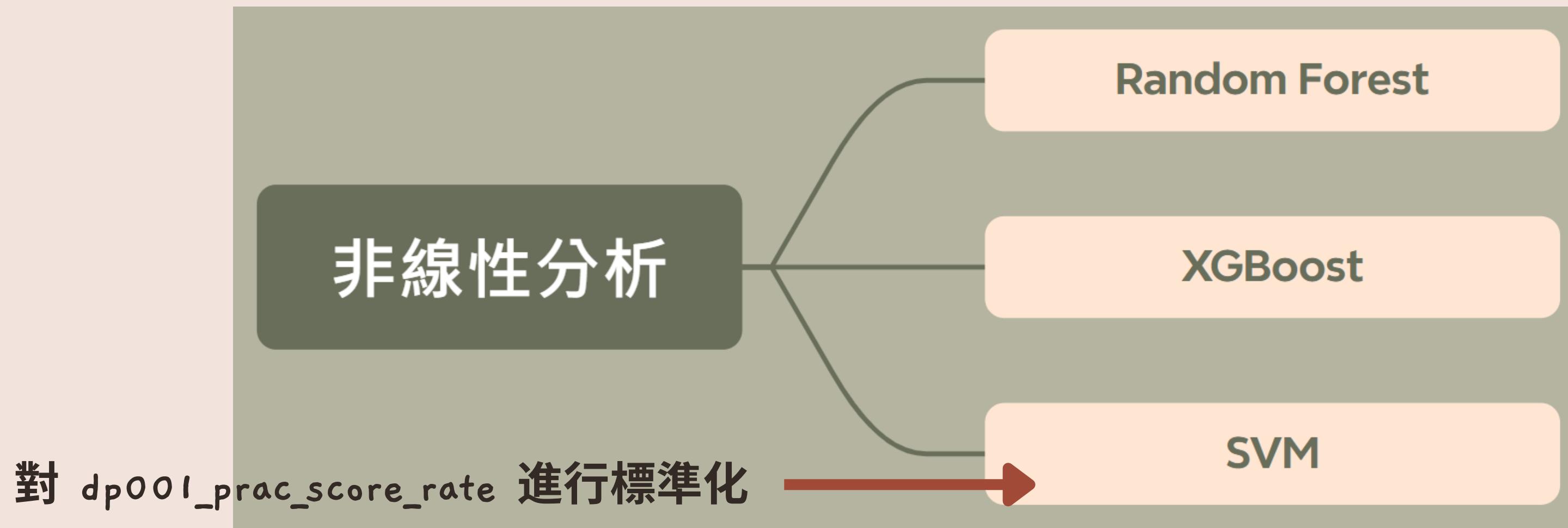
研究問題 2 練習分數與影片學習成效之關聯

對 `dp001_prac_score_rate` 進行標準化



研究問題 2

練習分數與影片學習成效之關聯



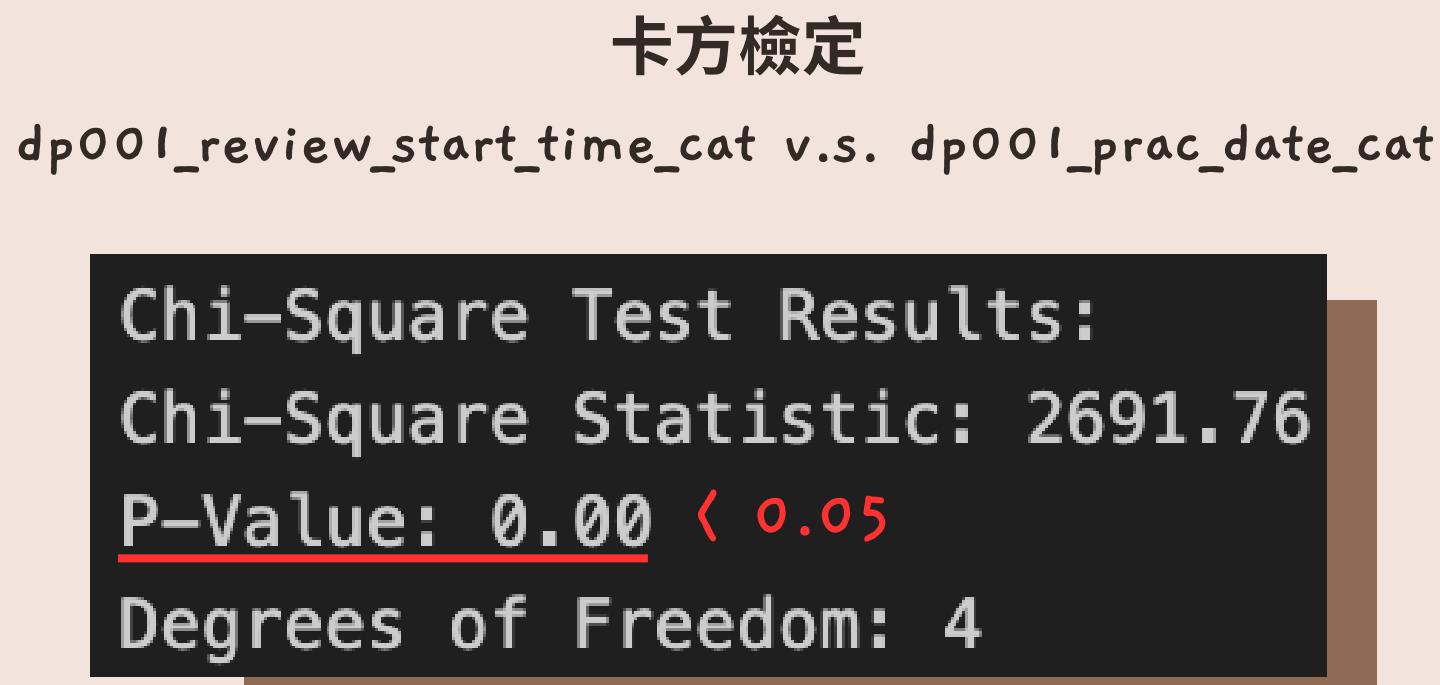
RESULT

研究問題 1

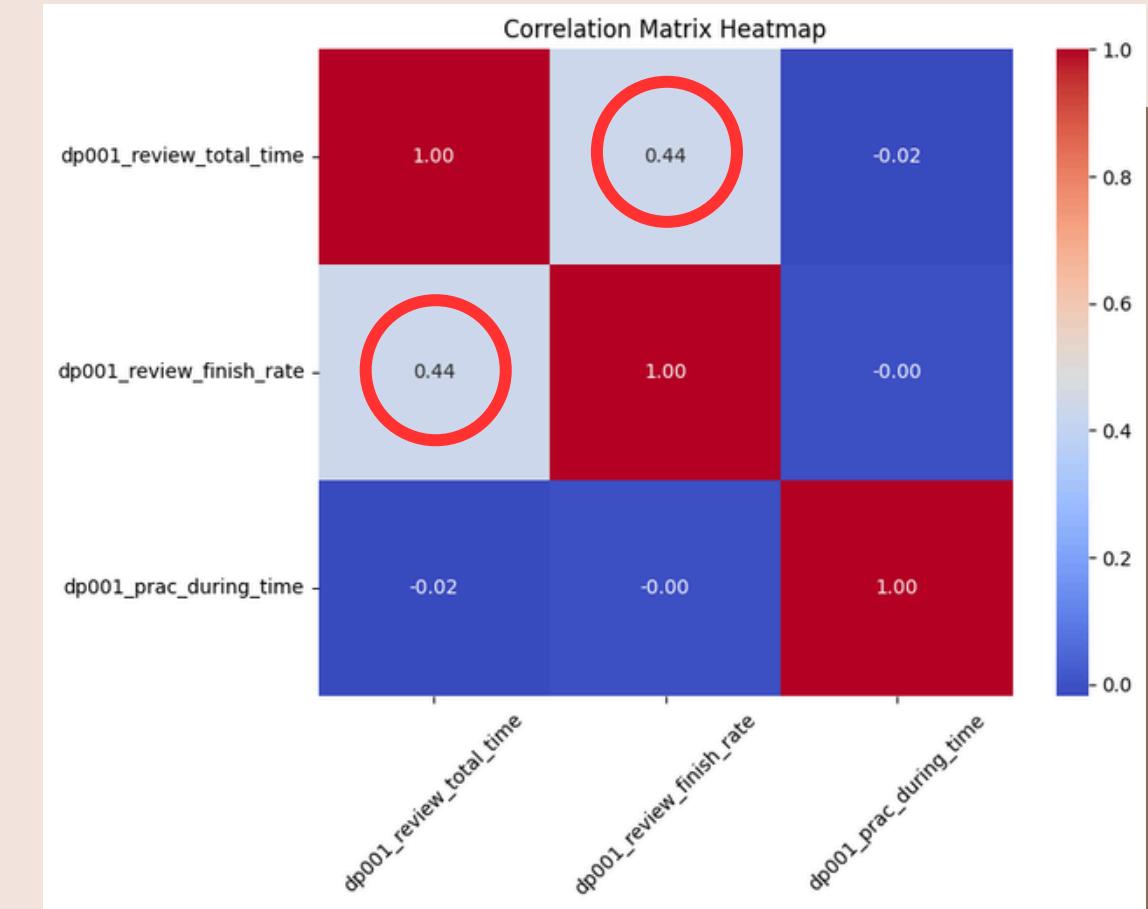
影片觀看行為、練習時長和學習時間點是否會影響練習分數？

1. 找出各個欄位(依變項)之間是否具有相關：

(a) 類別型依變項



(b) 連續型依變項



- 結果是顯著的，代表觀看時間與練習時間具有相關

- 影片觀看行為的變數之間呈中度相關 ($r=0.44$)
- 其餘變數間呈接近無相關

研究問題 1

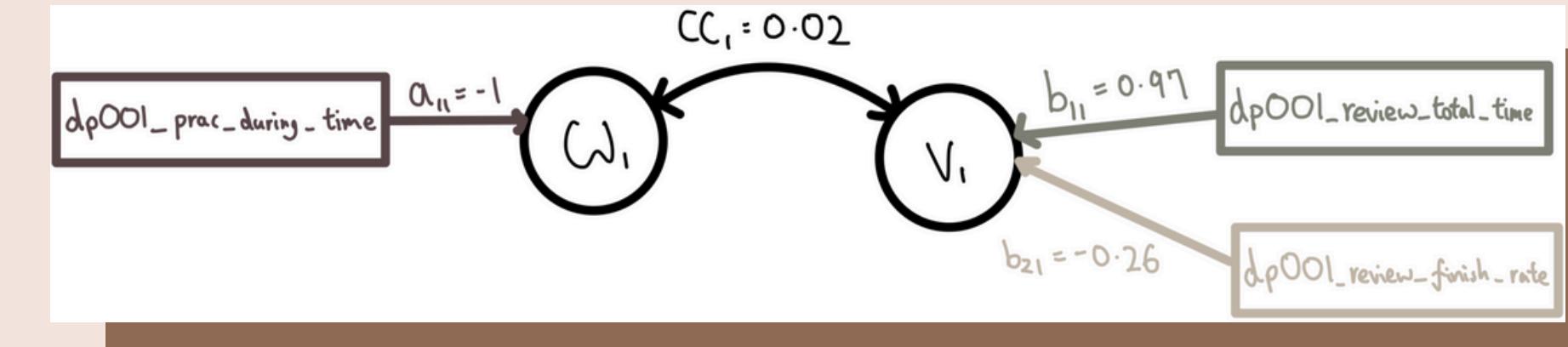
影片觀看行為、練習時長和學習時間點是否會影響練習分數？

2. 找出因素之間是否具有相關：

(a) 由連續型依變項所組成的因素 CCA

Canonical correlations:
[0.02]

Canonical Weights:
X :
[[0.9660081]
[-0.25851182]]
Y:
[[-1.]]



- 從 $CC1 = 0.02$ 可以得出，練習時長與影片觀看行為沒有相關

研究問題 1

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

2. 找出因素之間是否具有相關：

(b) 學習時間點 vs 影片觀看行為 vs 練習時長 MANOVA

```
Box's M test for dp001_prac_date
  Chi2    df   pval  equal_cov
box   inf  12.0   0.0     False
```

```
Box's M test for dp001_review_start_time
  Chi2    df   pval  equal_cov
box   inf  12.0   0.0     False
```

```
HZResults(hz=185.61024212594012, pval=8.437463739713922e-305, normal=False)
HZResults(hz=170.56207935251578, pval=2.45650683227867e-292, normal=False)
HZResults(hz=113.62643912220051, pval=3.273074844012646e-213, normal=False)
HZResults(hz=185.74663521136836, pval=2.482183879313664e-306, normal=False)
HZResults(hz=168.1360684399676, pval=7.905181407995131e-286, normal=False)
HZResults(hz=116.19166028743331, pval=3.9871299813687093e-219, normal=False)
```

Multivariate linear model						
	Intercept	Value	Num DF	Den DF	F Value	Pr > F
	Wilks' lambda	0.9978	3.0000	2372.0000	1.7510	0.1545
	Pillai's trace	0.0022	3.0000	2372.0000	1.7510	0.1545
	Hotelling-Lawley trace	0.0022	3.0000	2372.0000	1.7510	0.1545
	Roy's greatest root	0.0022	3.0000	2372.0000	1.7510	0.1545
	dp001_prac_date_cat	Value	Num DF	Den DF	F Value	Pr > F
	Wilks' lambda	0.9930	6.0000	4744.0000	2.7948	0.0103
	Pillai's trace	0.0070	6.0000	4746.0000	2.7915	0.0103
	Hotelling-Lawley trace	0.0071	6.0000	3160.8895	2.7986	0.0102
	Roy's greatest root	0.0069	3.0000	2373.0000	5.4454	0.0010
	dp001_review_start_time_cat	Value	Num DF	Den DF	F Value	Pr > F
	Wilks' lambda	0.9892	6.0000	4744.0000	4.3200	0.0002 < 0.001
...	Hotelling-Lawley trace	0.0106	12.0000	4146.6896	2.1010	0.0141
	Roy's greatest root	0.0077	4.0000	2374.0000	4.5695	0.0011

- Box's M test 結果顯著 ($p\text{-value} < 0.0001$)，不符合變異數同質性
- 資料不符合多元常態分佈
- MANOVA 結果顯示影片觀看行為和練習時長因素對於觀看時間之間存在顯著差異

研究問題 1

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

2. 找出因素之間是否具有相關：

(b) 學習時間點 vs 影片觀看行為 vs 練習時長 MANOVA

OLS Regression Results						
	coef	std err	t	P> t	[0.025	0.975]
Dep. Variable:	dp001_review_total_time	R-squared:	0.020			
Model:	OLS	Adj. R-squared:	0.017			
Method:	Least Squares	F-statistic:	6.162			
Date:	Tue, 24 Dec 2024	Prob (F-statistic):	6.78e-08			
Time:	00:23:32	Log-Likelihood:	-3356.8			
No. Observations:	2383	AIC:	6732.			
Df Residuals:	2374	BIC:	6784.			
Df Model:	8					
Covariance Type:	nonrobust					
Intercept	0.0663	0.037	1.786	0.074	-0.006	0.139
dp001_prac_date_cat[T.evening]	-0.3104	0.116	-2.670	0.008	-0.538	-0.082
dp001_prac_date_cat[T.morning]	-0.3003	0.128	-2.354	0.019	-0.551	-0.050
dp001_review_start_time_cat[T.evening]	0.4704	0.113	4.156	0.000	0.248	0.692
dp001_review_start_time_cat[T.morning]	-0.1321	0.101	-1.310	0.190	-0.330	0.066
dp001_prac_date_cat[T.evening]:dp001_review_start_time_cat[T.evening]	-0.2161	0.165	-1.310	0.190	-0.539	0.107
dp001_prac_date_cat[T.morning]:dp001_review_start_time_cat[T.evening]	0.0320	0.231	0.138	0.890	-0.422	0.486
dp001_prac_date_cat[T.evening]:dp001_review_start_time_cat[T.morning]	-0.4333	0.403	-1.074	0.283	-1.224	0.358
dp001_prac_date_cat[T.morning]:dp001_review_start_time_cat[T.morning]	0.2948	0.162	1.820	0.069	-0.023	0.612
Omnibus:	561.621	Durbin-Watson:	0.733			
...						

< 0.05

- 若與MANOVA的結果取交集，影片觀看行為和練習時長因素確實對於觀看時間之間存在顯著差異

研究問題 1

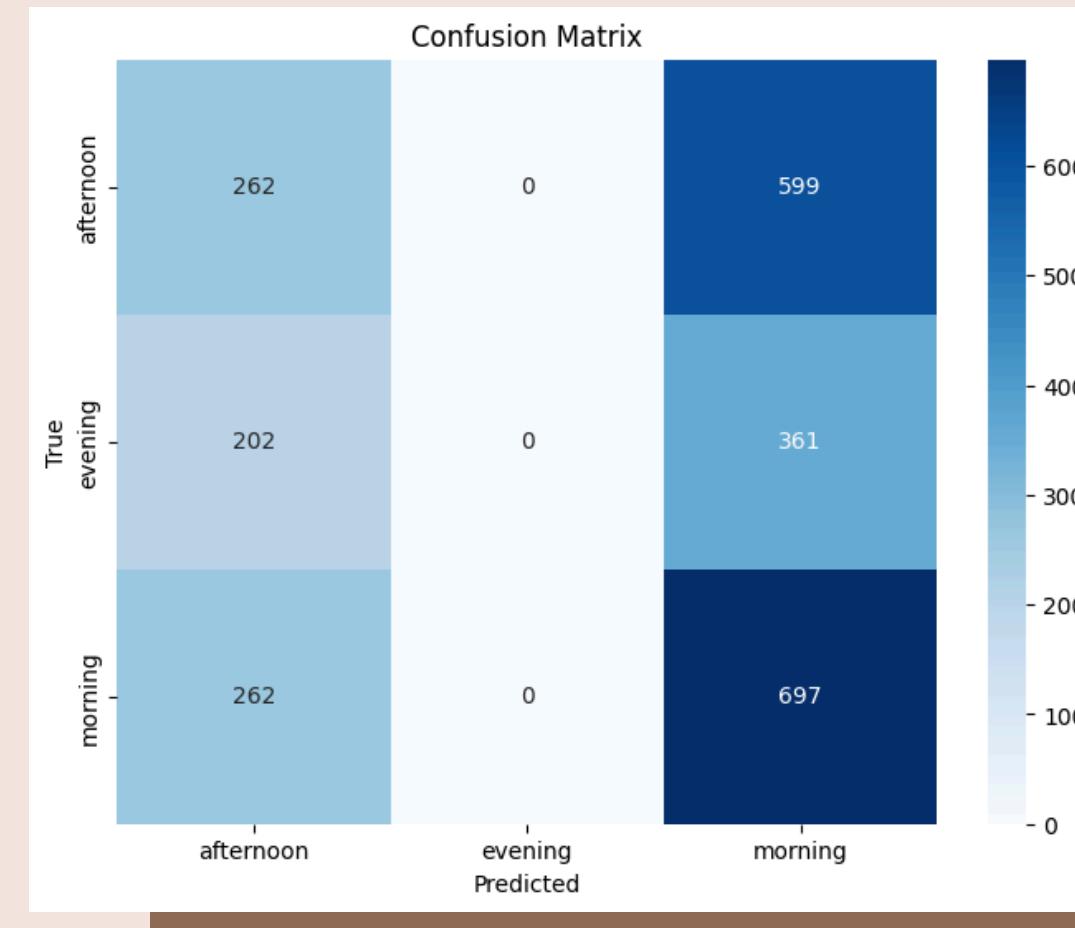
影片觀看行為、練習時長和學習時間點是否會影響練習分數？

2. 找出因素之間是否具有相關：

(b) 學習時間點 vs 影片觀看行為 vs 練習時長

針對觀看時間進行事後檢定 (DDA)

```
Canonical Scores:  
[[ 1.05 -0.45]  
 [ 1.05 -0.45]  
 [ 0.57 -0.01]  
 [ 0.57 -0.01]  
 [-2.27 -1.81]  
 [-2.27 -1.81]] ...  
  
Intercept of 3 canonical functions:  
[-1.02 -1.45 -0.91]  
  
Coefficients of 3 canonical functions:  
[[ 0.04 -0.06  0.04]  
 [ 0.07  0.08 -0.01]  
 [-0.08  0.    -0.02]]  
  
Explained Variance Ratio:  
[0.71  0.29]
```



```
Accuracy Rate:  
0.402  
  
Classification Report:  
precision recall f1-score support  
  
afternoon      0.36   0.30   0.33   861  
evening        0.00   0.00   0.00   563  
morning       0.42   0.73   0.53   959  
  
accuracy       0.40   0.40   0.40   2383  
macro avg      0.26   0.34   0.29   2383  
weighted avg   0.30   0.40   0.33   2383
```

- Discriminant function 1 可以解釋約 71% 的變異量，且總瀏覽時間與練習時長具有一樣的正向影響力；影片完成率具有負向影響力
- 模型的預測正確率僅 40%，並不能準確的預測觀看時間

研究問題 1

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

3. 找出因素與練習分數是否具有相關以及重要特徵：

(a) 類別型依變項

卡方檢定

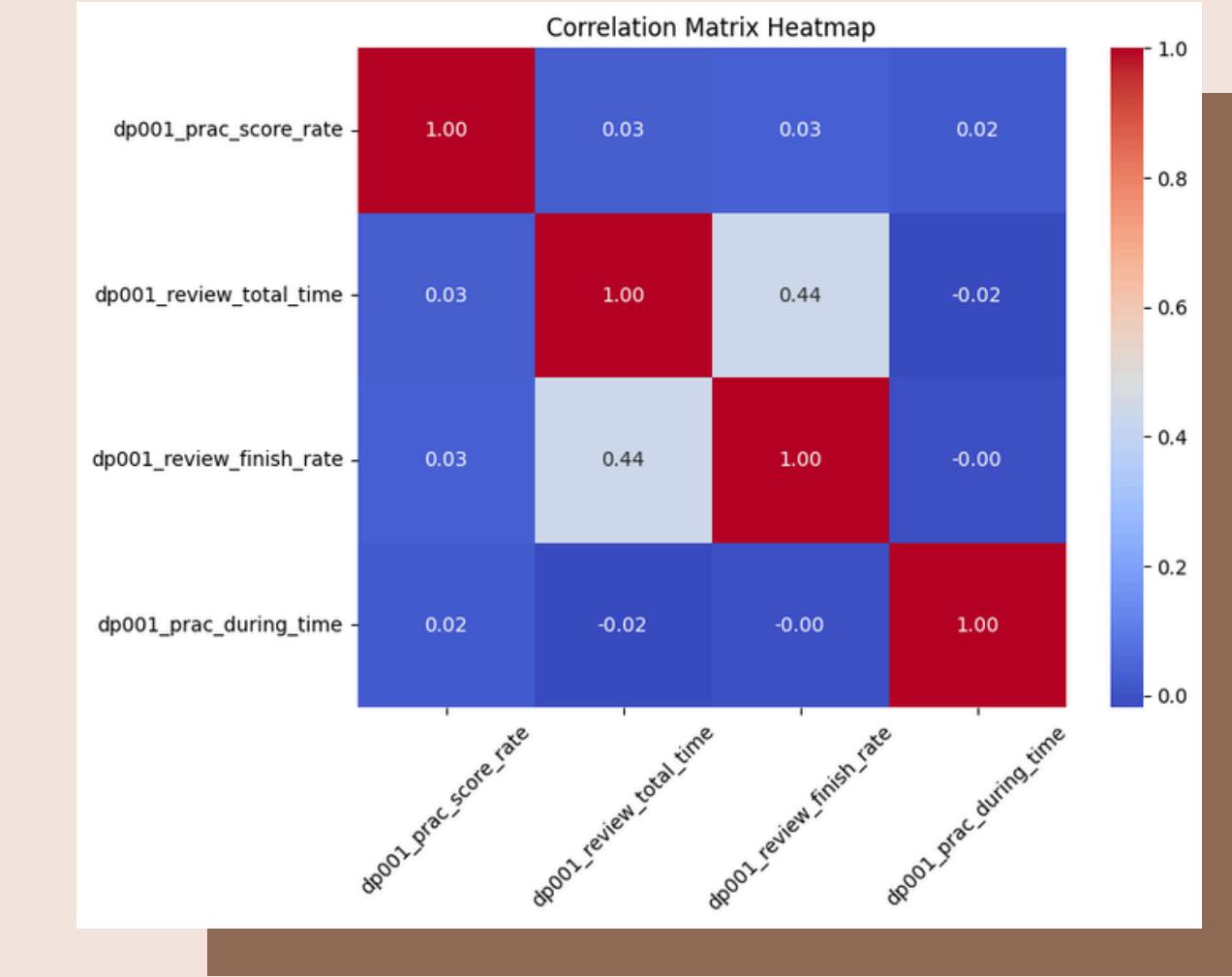
dp001_prac_date_cat

Chi-Square Test Results:
Chi-Square Statistic: 7.28
P-Value: 0.12
Degrees of Freedom: 4

dp001_review_start_time_cat

Chi-Square Test Results:
Chi-Square Statistic: 8.09
P-Value: 0.09
Degrees of Freedom: 4

(b) 連續型依變項



- 結果皆為不顯著，代表練習時間和觀看時間都與練習分數不具有相關性

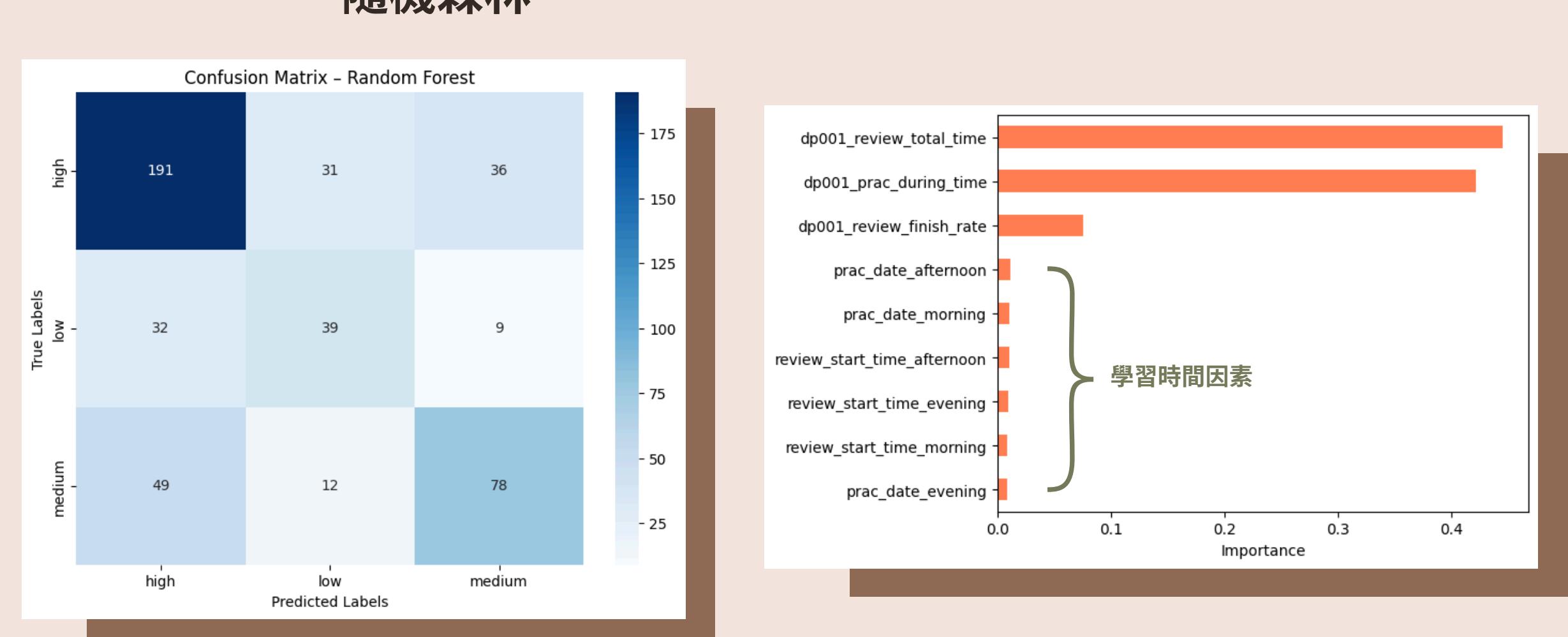
- 練習分數與連續型依變項們皆為接近無相關

研究問題 1

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

3. 找出因素與練習分數是否具有相關以及重要特徵：

Classification Report:				
	precision	recall	f1-score	support
high	0.70	0.74	0.72	258
low	0.48	0.49	0.48	80
medium	0.63	0.56	0.60	139
accuracy			0.65	477
macro avg	0.60	0.60	0.60	477
weighted avg	0.64	0.65	0.64	477



- 模型預測測試資料的準確率約65%，且練習分數為高的預測表現較好
- 重要的預測特徵：總瀏覽時間 > 練習時長 > 影片完成率

研究問題 2

練習分數與影片學習成效是否相關？

Logistic Regression：練習分數與學習成效的直接關係不強烈

	coef	std err	z	P> z	[0.025	0.975]

const	0.6357	0.043	14.686	0.000	0.551	0.721
dp001_prac_score_rate	0.2165	0.042	5.098	0.000	0.133	0.300

< 0.05

研究問題 2

練習分數與影片學習成效是否相關？

Sobel Test：練習分數非以下 3 個因子的中介變數



Sobel Test 結果：

Sobel Z-score: 1.38

P-value: 0.17 > 0.05

總瀏覽時間

Sobel Test 結果：

Sobel Z-score: 1.48

P-value: 0.14 > 0.05

影片完成率

Sobel Test 結果：

Sobel Z-score: 0.84

P-value: 0.4 > 0.05

練習時長

研究問題 2

練習分數與影片學習成效是否相關？

Interaction Analysis：學習行為變數皆未有效調節練習分數對學習成效的影響

	coef	std err	z	P> z
const	0.6362	0.043	14.687	0.000
dp001_prac_score_rate	0.2150	0.043	5.052	0.000
dp001_review_total_time	0.0496	0.044	1.134	0.257
interaction	-0.0049	0.045	-0.110	0.912

> 0.05

總瀏覽時間

研究問題 2

練習分數與影片學習成效是否相關？



Random Forest

Confusion Matrix:
[[28 115]
 [141 194]]

Classification Report:

	precision	recall	f1-score	support
0.0	0.17	0.20	0.18	143
1.0	0.63	0.58	0.60	335
accuracy			0.46	478
macro avg	0.40	0.39	0.39	478
weighted avg	0.49	0.46	0.48	478



XGBoost

Confusion Matrix:
[[28 115]
 [96 239]]

Classification Report:

	precision	recall	f1-score	support
0.0	0.0	0.23	0.20	143
1.0	1.0	0.68	0.71	335
accuracy			0.56	478
macro avg	0.45	0.45	0.45	478
weighted avg	0.54	0.56	0.55	478



SVM

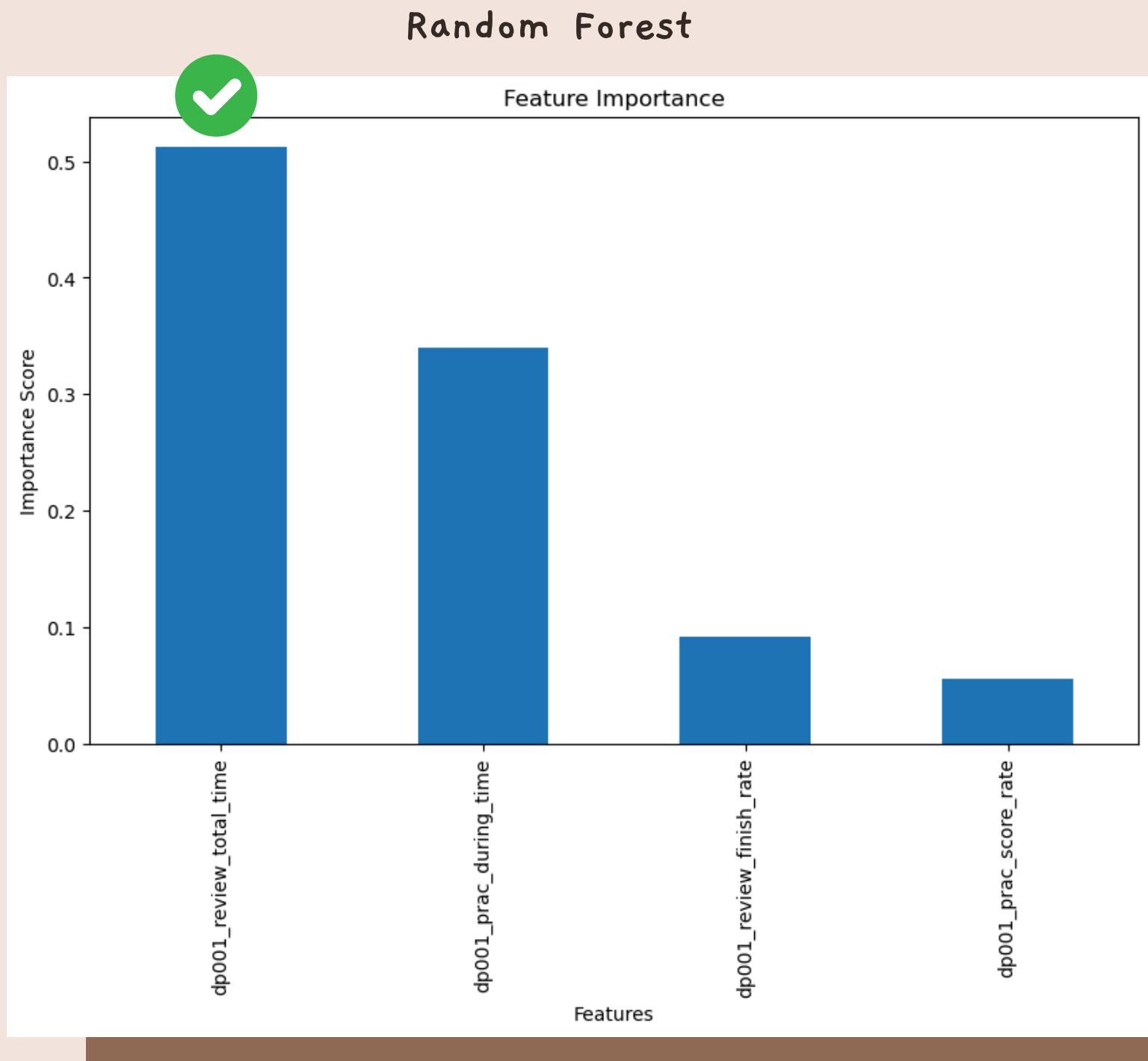
Confusion Matrix:
[[77 66]
 [136 199]]

Classification Report:

	precision	recall	f1-score	support
0.0	0.0	0.36	0.54	143
1.0	1.0	0.75	0.59	335
accuracy			0.58	478
macro avg	0.56	0.57	0.55	478
weighted avg	0.63	0.58	0.59	478

研究問題 2

練習分數與最終學習成效是否相關？



研究問題 2

練習分數與最終學習成效是否相關？

Linear SVM's weight

```
Weights of Linear SVM:  
✓ dp001_prac_score_rate: 0.0920  
dp001_review_finish_rate: 0.0360  
dp001_prac_during_time: 0.0105  
dp001_review_total_time: 0.0098
```

→ 標準化後得到的最高權重依然很小

CONCLUSION

RQ1 CONCLUSION

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

1. 三個因素之間沒有相關性，且均與練習分數無顯著關聯。

重要的預測特徵：

2. 總瀏覽時間 > 練習時長 > 影片完成率

RQ2 CONCLUSION

練習分數與最終學習成效是否相關？

1. 練習分數與學習成效之間的直接關係弱 ($\text{Corr} = 0.2165$)
2. 相比練習分數，總學習時間對學習成效更重要



結論

1. 練習分數無法有效反映學習行為的投入程度與影片學習成效
 2. 學習行為直接影響影片學習成效
 3. 總瀏覽時間是提升影片學習成效的關鍵
- 追蹤學習行為並優化時間管理，提升學習投入與成效。



FUTURE WORK

FUTURE WORK

1. 利用 SHAP 分析總瀏覽時間，揭示其對學習成效的影響程度與方向
2. 擴展分析變項，探索更多影響因素
3. 擴大研究對象與平台，驗證結果一致性



參考文獻

- Dancey, C. and Reidy, J. (2004) **Statistics without Maths for Psychology: using SPSS for Windows.** Prentice Hall, London.
- Akbulut, Y. (2024). **Impact of different practice testing methods on learning outcomes.** European Journal of Education. <https://doi.org/10.1111/ejed.12626>.
- 顏志龍。 (2017)。 中介效果分析的各種方法及其限制。 銘傳心理, 顏志龍的部落格。 <https://dragon1943.blogspot.com/2017/11/blog-post.html>

THANK YOU !



Q & A

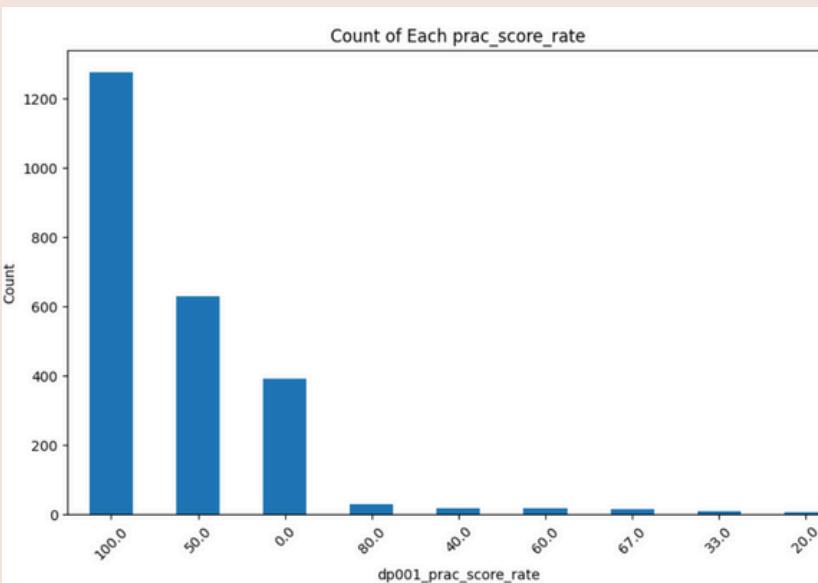
RQ I

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

1. 在練習分數的分組中，是根據什麼依據對三個組別進行切分？
2. 1:49，練習分數從40和70分去切高中低，為什麼會選這兩個分數點切呢？
3. 請問，為什麼選擇 40 和 70 作為分界點？是否是進行分位數切分或基於分布的自然分割點進行劃分之類的

Answer:

會將練習分數分成三組是因為練習分數的分佈從0~100，且有9個unique values，所以根據直覺與經驗，選擇分成三組。至於會從40和70分去切高中低是因為當初是先參考ChatGPT的建議（我們將練習分數的所有unique value喂給它，請它幫我們分成三組）。考慮到原本練習分數的分佈以及分組後的各個組別樣本數，我們認為這個分法是合理的。



dp001_prac_score_rate_cat	
high	1304
medium	677
low	406



RQ I

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

4. 請問凌晨的資料筆數約為多少？
 5. 想請問你們是否有額外嘗試過保留凌晨的資料進行分析？
 6. 想請教由於凌晨資料數過少，選擇將其資料點刪除，跟沒有刪除，會有甚麼差異嗎？
-

Answer:

只有練習時間有屬於凌晨（00:00~05:59）的資料，且只有4筆資料(約佔所有資料的0.17%)

dp001_prac_date_cat	
morning	944
afternoon	912
evening	527
early morning	4

一開始有保留凌晨的資料進行分析，但是沒辦法進行MANOVA，會出現 ValueError: Covariance of x singular! 的錯誤（如圖所示）。由於出現錯誤，我們才去檢查資料的分佈以及各組的資料數，進而發現凌晨的資料數過少。刪除凌晨的資料後就沒有出現錯誤了。



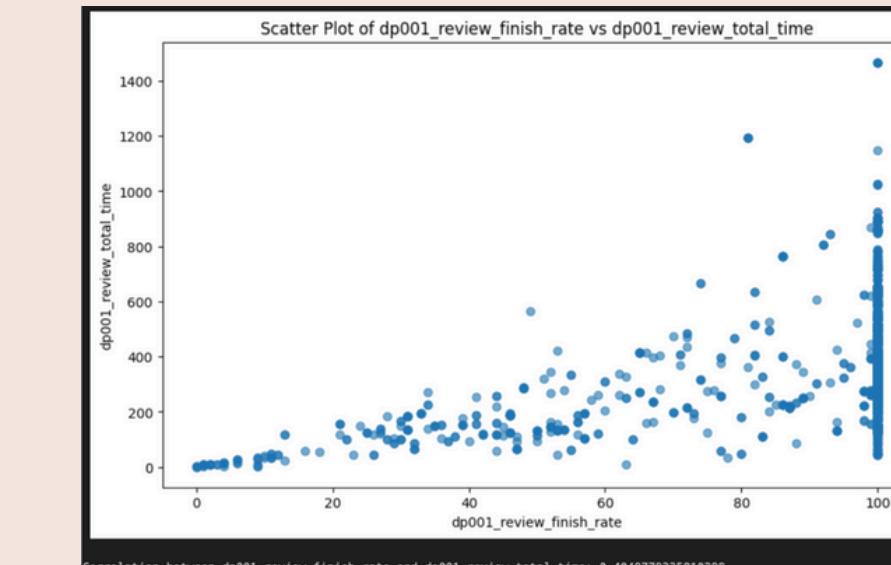
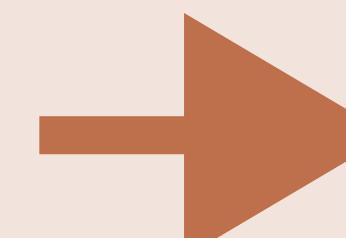
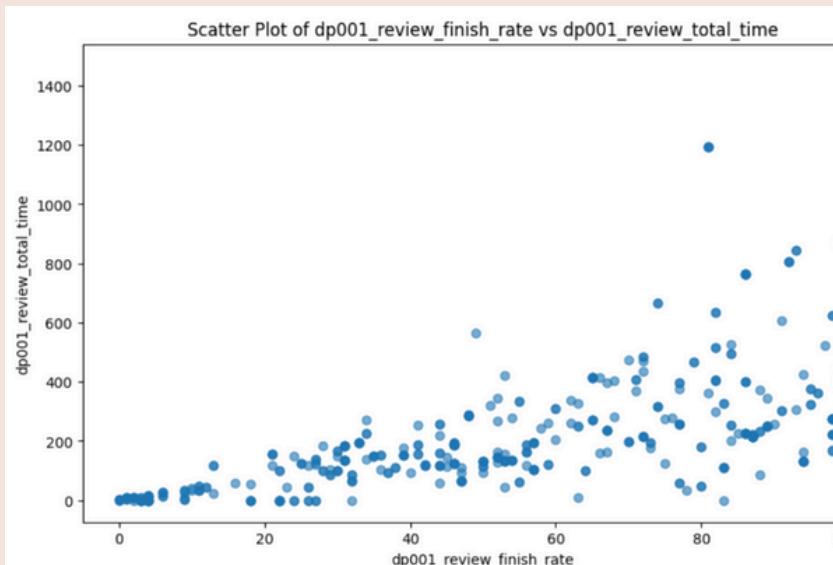
RQ I

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

7. 影片完成率與影片觀看時間呈中度相關，是否因為有掛網等極端值？請問請端值怎麼處理呢？

當初在分析的時候，我們沒有考慮到極端值（掛網的情況比較難從資料得知）。

現在重新檢查資料後，發現有影片完成率=100且總瀏覽時間=0的情況（共44筆），如果將他們刪除， $r= 0.40$ 小於刪除前的 $r=0.44$ ，所以影片完成率與影片觀看時間呈中度相關可能有受到極端值的影響，但是相關性仍為中度相關。



RQ I

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

8. 關於CCA分析，由 dp001_prac_during_time 所構成的 W 變項，兩者之間的係數為 -1，請問你們會如何解釋這個結果？

Answer:

典型權重為-1代表練習時長與其所構成的W變項呈完全負相關，代表任何練習時長的變化將對典型變量W產生最大可能的負向影響。



RQ I

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

9. 請問第15頁使用MANOVA前發現同質性跟常態性都不顯著但還是繼續使用MANOVA的原因是什麼 是否有考慮使用其他更穩健的檢定方法或是將變數做標準化等等
-

Answer:

參考ChatGPT，即使假設不完全滿足，MANOVA 在樣本量足夠大且組間樣本數大致平衡時，仍具有一定的穩健性。

沒有考慮使用其他更穩健的檢定方法，但是有將變數做標準化，之後會參考同學的建議，嘗試使用其他方法如 Robust MANOVA。



RQ I

影片觀看行為、練習時長和學習時間點是否會影響練習分數？

10. 變項處理1:40，將觀看時間分成早中晚(類別)、練習分數高中低(類別)，對應到Result 2:40，左半邊卡方檢定，請問是將觀看時間的早中晚和練習分數高中低做卡方嗎？簡報顯示其中一個變項是練習時間，練習時間屬於連續變項，不應用卡方(是要做練習分數高中低or練習時間哪一個?)，這裡有疑慮。
11. 第一個卡方分析那邊，是指將觀看影片的時間跟做練習題的時間都分成早中晚之後檢驗兩者之間是否有關聯嗎？分析結果是有關聯的，這樣可以怎麼解釋這個結果呢？

Answer:



不是，這裡是要找出各個欄位(依變項)之間是否具有相關，所以是將觀看時間的早中晚和練習時間的早中晚做卡方檢定。另外，練習時間應該是屬於類別變項 (原本：2023/6/29 15:48:08，轉換後：afternoon)。

觀看時間和練習時間有關聯的，代表這兩個變項之間並非獨立的，且顯示學生在學習的時間安排具有某種關聯性，所以可以視為同一個因素「學習時間點」。



RQ2 練習分數與影片學習成效是否相關？

1. 在交互作用分析中，結果顯示「總瀏覽時間」對學習成效的影響不顯著；然而，在機器學習模型中，「總瀏覽時間」卻是學習成效的重要預測特徵。請問你們怎麼看待這兩個結果之間的差異？
-

Answer:

我認為這表示「總瀏覽時間」的影響可能難以用線性模型來捕捉，這可能是因為其與學習成效之間存在非線性關係或交互作用，而在機器學習模型中「總瀏覽時間」是學習成效的重要預測特徵，意味著「總瀏覽時間」可能並非單獨直接影響學習成效，而是與其他變數共同作用，形成更複雜的影響模式



RQ2 練習分數與影片學習成效是否相關？

1. 在交互作用分析中，結果顯示「總瀏覽時間」對學習成效的影響不顯著；然而，在機器學習模型中，「總瀏覽時間」卻是學習成效的重要預測特徵。請問你們怎麼看待這兩個結果之間的差異？
-

Answer:

我認為這表示「總瀏覽時間」的影響可能難以用線性模型來捕捉，這可能是因為其與學習成效之間存在非線性關係或交互作用，而在機器學習模型中「總瀏覽時間」是學習成效的重要預測特徵，意味著「總瀏覽時間」可能並非單獨直接影響學習成效，而是與其他變數共同作用，形成更複雜的影響模式



RQ2 練習分數與影片學習成效是否相關？

2. 選用XGboost而非ANN的原因為何？XGboost有什麼特點？

Answer:

ANN 包含多層神經元和大量參數，性能依賴於大量的訓練數據，因為其模型，需要足夠的數據來避免 overfitting，而XGBoost 是一種基於決策樹的方法，對資料量的需求相對較低。



RQ2 練習分數與影片學習成效是否相關？

3. 從報告中得知SVM的accuracy最高但無法成功預測y所以被排除使用，但我好奇若繼續使用SVM來選取重要特徵，其情況是否與RF或XGboost相似，是否也能驗證total_time是重要特徵？

Answer:

因為非線性核（如 RBF）會將資料映射到高維特徵空間，使特徵的重要性無法透過權重提取，只能透過某些技巧來間接地協助特徵選擇，這邊使用權重向量。

重要特徵名稱與權重：

```
dp001_prac_score_rate: 0.0025
dp001_review_total_time: 0.0
dp001_review_finish_rate: 0.0017
dp001_prac_during_time: 0.0
```

當使用線性核（linear kernel）時，SVM 會產生一個權重向量 w ，這個向量表示每個特徵對分類的影響程度。當該權重絕對值越大，則對分類邊界的貢獻也越大。可以看到相對較高的是練習分數 & 影片完成率，他們對模型預測 $y=1$ 有輕微的正向影響，但這邊的結果就跟RF或XGboost不同。



RQ2 練習分數與影片學習成效是否相關？

- dp001_exam_binary_res 是二元的資料型態，你們將他當作依變項的話，是如何處理的呢？
-

Answer:

在線性分析的部分，我們採用的是 Logistic Regression（邏輯回歸），主要用於二元分類，中介變數和調節變數檢驗也是結合了邏輯回歸來進行的
在 RF、XGB 和 SVM 的部分，則是採用分類模型來預測其為 0 或 1



綜合問題

1. 在進行機器學習分析時，你們採用了什麼樣的方式來訓練和測試資料集？
-

Answer:

我們都是將資料集分成**80%**的訓練集和**20%**的測試集。

```
# Splitting into training and testing sets  
train, test = train_test_split(rf1_data, test_size=0.2, random_state=2024)
```



綜合問題

2. dp001的exam是在觀看影片的過程中會出現的檢核點題目，prac則是獨立於影片的練習題，這樣的話是不是將prac當作最終學習成效是比較合理的呢？
-

Answer:

我們認為將 exam 作為學習成效的主要依據，是因為 exam 直接嵌入影片學習過程中，能即時檢核學生對影片內容的理解，與我們研究的影片學習情境高度匹配。因此，基於評估時間點與干擾因素考量，使用 **exam** 更能準確反映影片學習的成效。

基於上述論點，我們認為可以重新定義 exam 變數為影片學習成效



綜合問題

3. 極端值影響除了影片觀看時間的極端值，其他變數（如檢核點答題正確率和練習題正確率）是否也存在異常值？在進行統計分析（例如ANOVA或CCA）之前，是否有進行極端值的篩選或替換？

Answer:

```
---最終學習成效---
Value counts:
dp001_exam_binary_res
1.0    1557
0.0     830
Name: count, dtype: int64
---練習分數---
Value counts:
dp001_prac_score_rate
100.0    1276
50.0      629
0.0       392
80.0      28
40.0      18
60.0      17
67.0      13
33.0       9
20.0       5
```

在進行統計分析前，沒有進行極端值的篩選，但是有將資料進行標準化。



綜合問題

4. 將影片觀看時間分為早、中、晚三個時段，這是一個非常有趣的思路。最後的結論中，7:20有強調「總瀏覽時間是提升學習成效的關鍵」，並提到「追蹤學習行為並優化時間管理」。然而，學生一天在使用線上平台學習的時間是有限的，因此是否有研究或建議，針對「在哪個時段（早、中、晚）投入較長時間學習，能夠最有效地提升學習成效」這一方面進行探討呢？
-

Answer:

因為時間有限，我們尚未能針對「在哪個時段（早、中、晚）投入較長時間學習，能夠最有效提升學習成效」進行探討，但對於此問題，討論後我們認為可以採用：

1. 根據學生在早、中、晚的瀏覽時間進行分組，計算各時段的學習成效均值，使用統計檢定（如 ANOVA）比較不同時段的成效差異
2. 建立回歸模型，將早、中、晚的瀏覽時間作為自變數，學習成效作為因變數，分析各時段學習時間對成效的影響力



綜合問題

老師提問：有文獻，框架明確易懂（跟RQ1&RQ2是不是相反了？）。使用多種分析方法(XGboost!)，分析假設檢驗完整，結果呈現很清楚。迴歸分析有檢驗標準化係數嗎？SHAP是？

- RQ1是針對學習行為（影片觀看行為、練習時長和學習時間點）跟練習分數間的關係，而RQ2是研究「練習分數」與「學習成效」之間的關聯。
- 在進行迴歸分析時，尚未對變數進行標準化處理，這部分會再補上！
- SHAP 用於解釋機器學習模型預測結果的方法，透過分配各個特徵對預測結果的貢獻值，幫助我們清楚知道「特徵值高低會如何影響預測值」。**SHAP 值 > 0** 表示該特徵增加預測值



綜合問題

2. 資料檔1:16，用什麼共同變項進行data1和data2資料合併呢？是PseudoID嗎？
 3. 報告中提到將Data1 和 Data2 進行合併，並成功生成 2387 筆合併數據。請問，合併這兩個數據集的具體方法是什麼？
 4. 能將兩筆資料進行合併很厲害，因為我們這組合併時一直發生錯誤最後放棄，所以想請問你們是用什麼方法順利將資料合併的？
-

Answer:

不是，因為當初無法確認data1和data2的“PseudoID”是否相同（同樣PseudoID為同一位使用者），所以我們調整data1的“PseudoID”的編碼後，把data1的資料加入data2

經過老師證實，同樣PseudoID為同一位使用者，因此我們會再以樣PseudoID作為共同變項進行data1和data2資料合併

```
# Get the column order from Data Frame 2
data2_columns = data2.columns

# Reorder Data Frame 1 columns to match Data Frame 2
data1 = data1.reindex(columns=data2_columns)

# Add 405 to the "PseudoID" column values
data1.iloc[:, 0] = data1.iloc[:, 0] + 405

# Concatenate data1 and data2 along the rows (axis=0)
bigdata = pd.concat([data2, data1], ignore_index=True)
```



綜合問題

5. 合併後1:22，合併後資料數有2387筆，請問有幾位學生？

```
num_users = q1_data['PseudoID'].unique().shape[0]
print(f"Number of unique users: {num_users}")
✓ 0.0s
Number of unique users: 258
```

