學習單1-數據分析

金企四甲 林嘉柔 409411167

- · 從 SRDA 網站上下載"教育相關"的資料,資料變數須包含名目/類別變數、連續變數,並將資料在 EXCEL 上建立一個 csv 檔。
- 針對此數據執行以下三項任務:
- Q1 執行一些統計敘述的分析,並說明其內容與您的發現。
- Q2 產出各種一維列聯表和二維列聯表分析,並說明其內容與您的發現。
- Q3 繪製講義內所學的視覺化圖形,並說明其內容與您的發現。
- **資料主題:**「實習學生的網路社會資本、求助行為、自我調整學習與教師信念的關聯之研究」。

一、問題描述:

對於進行教育實習的學生來說,雖然學校安排了輔導教師給他們並提供完善的輔導制度,但卻仍然面臨著在陌生環境中與輔導教師互動、班級經營、溝通困難和人際關係等問題。而先前研究指出,他們即便尋求輔導教師或其他人員的協助,也未必能得到充分的解決,因此,為因應這些困境,實習學生往往會透過網路社群及其他社會資本來尋求幫助和慰藉。

尋求幫助後,實習學生需要透過自我調整,根據不同的問題情境與回應調整實習任務,以因應教育專業的需要。而且在這些過程中,除了應用所學知識與技能外,他們也可能對教師的工作有重新地瞭解。其藉由不斷地與人互動、解決內在衝突並調整對此工作的看法,若其新的觀點能獲得他人認同,便可能塑造正向的教師信念。

然而,現今對於實習學生如何運用網路社會資本與此求助行為如何影響他們的自我調整學習策略,以及這些對教師信念的影響尚缺乏深入探討。因此,此教育資料主要透過深入了解他們在教育實習期間的行為和思考來進行研究,希望為提升教師培訓過程中的支援措施提供資訊。

參考資料:檔案連結 (研究動機與目的)

→→而此份學習單主要透過執行統計敘述分析、產出各種一維列聯表和二維列聯表,最後繪製視覺化圖形來更直接地觀察與佐證,進一步說明其內容與發現。

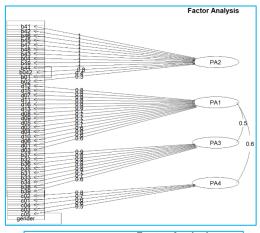
二、數據分析:

1. 資料說明與處理

變數多為五點量表 1~5 數值,代表非常不同意至非常同意,而資料報告中已透過信效度分析得知題項 c06、d14、e02 的相關性不顯著(因 r<0.05,採 0.95 的信心水準與 5%的信賴區間)因此先將此 3 題從資料中刪除。也在單因子變異數分析中得出無論實習生參與哪一個屬性學校的教育實習,在探討變項上沒有存在差異性,因此也將基本資料中觀察實習學校屬性的題項 school 刪除。

此外,題項 b0101~b0104 為將勾選「尋求協助時會聯繫的對象」變成類別型變數(勾選為1,無勾選為0),為了一併探討「聯繫對象」的重要性與對教師信念的關聯,以及方便進行整體的數據分析,因此將0、1 中勾選家人的1 改為1、大學老師的1 改為2、學長姐或同學(未在同一學校實習)為3、朋友為4,而因為題項可複選,所以改取平均值歸類至同一個變數 b01。(先以數值大小概略分析,看聯絡對象是否會有顯著影響,如果此項有顯著影響也可以再重新區分開來做為未來研究的方向之一。)

最後,在第4份學習單做完<u>因素分析進行維度縮減</u>後,刪除了幾項與 其對應的因素間相關性小於 0.5 的變數 ("gender"、" b01"、" b02"、"e05"與 "e09"),並將變數分類至各因素中,如下圖所示,而 各因素所代表的涵義由下表呈現:



自變數	因素涵義
d開頭	PA1→ "自我調整學習"
b4 開頭	PA2→ "網路社會資本中
	任務導向型相關"
b3 開頭	PA3→ "網路社會資本中
	情感連結型相關"
c開頭	PA4→ "求助動機"

	Factor Analysis			
e01 e03 e06 e12 e08 e15 e14 e13 e11 e07 e10	0.7 0.7 0.6 0.5 0.5 0.7 0.7 0.7 0.7 0.7 0.7 0.7 0.4 PA3 0.3 0.4			
e09 e04 e05	0.5 0.3			

應變數	因素涵義
e01, e03,	PA1→"教師與學生
e06, e08, e12	間之教學模式"
e11, e13,	PA2→ "校務參與"
e14, e15	
e07, e10	PA3→"教師與學生
	間自治與自省"
e04	PA4→ "教學目標"

而考量到測量的變數過多(自變數 40 個、應變數 12 個),為了方便分析與 能有更直觀的解釋,因此決定將自變數與應變數以其各四個因素中所涵蓋的各 個變數取平均值來代表,以此執行統計敘述分析、產出各種一維列聯表和二維 列聯表,以及繪製視覺化圖形,來觀察並說明其內容與發現。整理過後的資料 如下所示:

原X變數名稱	取平均後名稱	原Y變數名稱	取平均後名稱
b31~b39	b3	e01, e03, e06, e08, e12	e1
b41~b49	b4	e11, e13, e14, e15	e2
c01~c05	С	e07, e10	e3
d01~d16	d	e04	e4

2. 統計敘述分析與說明

```
> #統計運算
> average <- mean(table$b3, na.rm = TRUE)
> cat("b3平均數:", average, "\n")
b3平均數: 4.188913
> average <- mean(table$b4, na.rm = TRUE)
> cat("b4平均數:", average, "\n")
b4平均數: 3.188486
> average <- mean(table$c, na.rm = TRUE)
> cat("c平均數:", average, "\n")
c平均數: 4.130917
> average <- mean(table$d, na.rm = TRUE)
> cat("d平均數:", average, "\n")
d平均數: 3.98145
> average <- mean(table$e1, na.rm = TRUE)
> cat("e1平均數:", average, "\n")
e1平均數: 4.454584
> average <- mean(table$e2, na.rm = TRUE)
> cat("e2平均數:", average, "\n")
e2平均數: 4.273561
> average <- mean(table$e3, na.rm = TRUE)
> cat("e3平均數:", average, "\n")
e3平均數: 4.51919
> average <- mean(table$e4, na.rm = TRUE)
> cat("e4平均數:", average, "\n")
e3平均數: 4.51919
> average <- mean(table$e4, na.rm = TRUE)
> cat("e4平均數:", average, "\n")
e3平均數: 4.51919
```

從左圖的平均數來看可以發現,自變數的部分以 b3 最高,為 "網路社會資本中情感連結型相關",可知大多數人對其有較正面的想法與正向成長;而 b4 平均值最低,可能原因是有些人並無加入社群因此無填答;第二低的是 d "自我調整學習"面向,推測使用社交網路對外尋求協助,對實習學生的自我調整學習、自我認知等幫助略低,但整體來說還算不錯,接近"4"。

而應變數 "教師信念"的部分,以 e3 "教師與學生間自治與自省"的平均數最高,e2 "校務參與"的平均數最低,可知在實習學生運用網路社會資本與此求助行為,影響其自我調整學習策略,進而影響他們對 "教師信念"的看法中,又以教師應鼓勵學生自治與自動自發,以及教師應具備自我檢討能力最受認同;且對於教師應參與學校行政事務、規劃活動等相關事務的平均認同度最低。

而右圖的標準差中,雖然皆在 0.48 至 0.69 之間 (b4 因部分人無填答而不同), 整體來說皆小於 1,但在資料為 1~5 的數值 中,我認為整體集中度只能算是中偏高,若 標準差再小一些會更好。

```
sd_value <- sd(table$b3
cat("b3標準差:", sd_valu
                           sd_value,
b3標準差: 0.6263146
 > sd_value <- sd(table$b3, na.rm
> cat("b3標準差:", sd_value, "\n")
                                            "\n")
b3標準差: 0.6263146
> sd_value <- sd(table$b4, na.rm = TRUE)
> cat("b4標準差: 1.712384
 > sd_value <- sd(table$c, na.rm = TRUE)
> cat("c標準差:", sd_value, "\n")
|> cat("c標準差:", sd_value, "\n")
|c標準差: 0.6339432
|sed_value
> sd_value <- sd(table$e1, na.rm = TRUE)
> cat("e1標準差:", sd_value, "\n")
e1標準差: 0.4875428
> sd_value <- sd(table$e2, na.rm = TRUE)
> cat("e2標準差:", sd_value, "\n")
- ここの lese2, na.rm
- ここに e2標準差:", sd_value, "\n")
e2標準差: 0.5410345
> sd_value
 > sd_value <- sd(table$e3, na.rm = TRUE)
> cat("e3標準差:", sd_value, "\n")
> cat("e3標準差:", sd_value,
e3標準差: 0.5387293
  · sd_value <- sd(table$e4, na.rm = TRUE)
· cat("e4標準差:", sd_value, "\n")
e4標準差: 0.6873865
```

```
> cor(table$b3,table$e1) > cor(table$c,table$e1)
 [1] 0.2021174
                           [1] 0.1881032
 > cor(table$b3,table$e2) | > cor(table$c,table$e2)
                           [1] 0.2220535
 [1] 0.1774596
 > cor(table$b3,table$e3) ||> cor(table$c,table$e3)
                          [1] 0.1815753
 [1] 0.1963133
  cor(table$b3,table$e4)
                           cor(table$c,table$e4)
                          [1] 0.1643549
 [1] 0.1587384
   cor(table$b4, table$e1) ||> cor(table$d, table$e1)
[1] 0.3316352
                          [1] 0.207321
 > cor(table$b4,table$e2)
                           > cor(table$d, table$e2)
                         [1] 0.2914955
 [1] 0.2436149
 > cor(table$b4,table$e3)
                          > cor(table$d,table$e3)
 [1] 0.2260728
                           [1] 0.2094284
  cor(table$b4,table$e4)
                          > cor(table$d,table$e4)
 [1] 0.1884049
                          [1] 0.1510591
```

上圖為 X 變數與 Y 變數各構面中兩兩之間的相關性,雖然整體來說較低了點,但可看出皆為正相關,且由 b4 與 e1、d 與 e2 的相關性較高,可知 "網路社會資本中任務導向型相關"對"教師與學生間之教學模式"與"自我調整學習"對"校務參與"的影響較顯著。

3. 列聯表與說明

■ 各變數之一維列聯表

```
> #一維列聯表分析
> table(data$b3)
1 1.4 1.6 1.7 2 2.8 3 3.1 3.2 3.3 3.4 3.6 3.7 3.8 3.9 4 4.1 4.2 4.3 4.4 4.6 4.7 4.8 4.9 3 2 1 3 1 1 2 5 1 5 9 14 25 27 33 66 38 25 27 30 25 29 32 15 > prop.table(table(data$b3))
                                                   1.7
                                                                               2.8
                                     1.6
0.006396588 0.004264392 0.002132196 0.006396588 0.002132196 0.002132196 0.004264392 0.010660981
         3.2
                                                                 3.7
                       3.3
                                     3.4
                                                   3.6
                                                                               3.8
                                                                                             3.9
0.002132196 0.010660981 0.019189765 0.029850746 0.053304904 0.057569296 0.070362473 0.140724947
                       4.2
                                     4.3
                                                   4.4
                                                                 4.6
         4.1
                                                                               4.7
  .081023454 0.053304904 0.057569296 0.063965885 0.053304904 0.061833689 0.068230277 0.031982942
0.106609808
```

由上圖可得知在 b3 "網路社會資本中情感連結型相關"中,以4分最多,次多的為5分,大部分人皆為同意至非常同意,但仍有少數人填答的分數低於平均值。

```
> table(data$b4)
                     0 1.6 1.8 2 2.2 2.5 2.6 2.7 2.8 2.9
95 1 1 2 1
4.5 4.6 4.7 4.8 4.9
               1 3
1.9 5
                                        5 29 3 16 12 13 21 19 19 15 16 15 41
13 32 11 36 15 5
> prop.table(table(data$b4))
                  1.6
                                                  2.2
                                                             2.5
0.202558635 \ 0.002132196 \ 0.002132196 \ 0.004264392 \ 0.002132196 \ 0.006396588 \ 0.006396588 \ 0.008528785
       2.8
                 2.9
                                       3.1
                                                  3.2
                                                             3.3
                                                                       3.4
                                                                                  3.5
0.004264392 0.010660981 0.014925373 0.010660981 0.061833689 0.006396588 0.034115139 0.025586354
       3 6
                 3 7
                            3 8
                                       3 9
                                                   4
                                                            4 1
                                                                        4.2
                                                                                   4 3
0.027718550 0.044776119 0.040511727 0.040511727 0.031982942 0.034115139 0.031982942 0.087420043
                                                  4.8
       4.4
                 4.5
                            4.6
                                       4.7/
                                                            4.9
0.019189765 0.027718550 0.068230277 0.023454158 0.076759062 0.031982942 0.010660981
```

由上圖可得知在 b4 "網路社會資本中任務導向型相關"中,以 4.3 分最多,次多的為 4.8 分,整體來說較偏普通與同意。

```
> table(data$c)

1 1.6 1.8 2 2.4 2.8 3 3.2 3.4 3.6 3.8 4 4.2 4.4 4.6 4.8 5
1 2 1 4 1 6 14 9 19 28 45 115 47 37 48 24 68

> prop. table(table(data$c))

1 1.6 1.8 2 2.4 2.8 3 3.2

0.002132196 0.004264392 0.002132196 0.008528785 0.002132196 0.012793177 0.029850746 0.019189765
3.4 3.6 3.8 4 4.2 4.4 4.6 4.8

0.040511727 0.059701493 0.095948827 0.245202559 0.100213220 0.078891258 0.102345416 0.051172708
5
0.144989339
```

由上圖可得知在 C "求助動機"中,4 分的比例來到了 0.245,將近 4 分之 1,整體偏同意至非常同意。

```
> table(data$d)
1 1.6 1.8 2.2 2.3 2.5 2.6 2.7 2.8 2.9 3 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 4 4.1 4.2 4.3 4.4 4.5 1 1 2 1 1 1 1 3 1 3 11 10 10 10 6 18 10 36 26 60 50 45 24 31 12 29 4.6 4.7 4.8 4.9 5 13 20 11 10 12
> prop.table(table(data$d))
                                         1.8
                                                         2.2
                                                                        2.3
0.002132196 \ 0.002132196 \ 0.004264392 \ 0.002132196 \ 0.002132196 \ 0.002132196 \ 0.002132196 \ 0.002132196
                                                         3.1
          2.8
                          2.9
                                                                        3.2
                                                                                        3.3
0.002132196\ 0.006396588\ 0.023454158\ 0.021\underline{321962}\ 0.021\underline{321962}\ 0.021\underline{321962}\ 0.021\underline{321}962\ 0.012793177\ 0.038379531
                                                         3.9
          3.6
                          3.7
                                         3.8
                                                                                        4.1
                                                                                                        4.2
0.021321962 0.076759062 0.055437100 0 127931770 0.106609808 0.095948827 0.051172708 0.066098081
                                         4.6
                                                         4.7
0.025586354 0.061833689 0.027718550 0.042643923 0.023454158 0.021321962 0.025586354
```

由上圖可得知在 d"自我調整學習"中,整體分數多為 4 分同意。

由上圖可得知在 el "教師與學生間之教學模式"中,整體多為同意至 非常同意。

```
> table(data$e2)

1 1.8 2.5 2.8 3 3.3 3.5 3.8 4 4.3 4.5 4.8 5
1 1 2 1 8 14 22 46 89 87 76 52 70

> prop.table(table(data$e2))

1 1.8 2.5 2.8 3 3.3 3.5 3.8

0.002132196 0.002132196 0.004264392 0.002132196 0.017057569 0.029850746 0.046908316 0.098081023
4 4.3 4.5 4.8 5

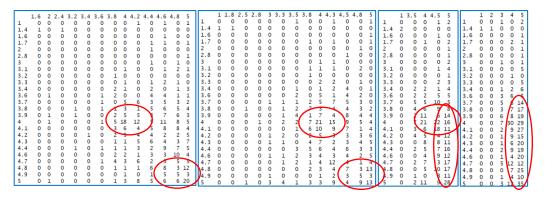
0.189765458 0.185501066 0.162046908 0.110874200 0.149253731
```

由上圖可得知在 e2 "校務參與"中,整體偏向同意。

由上圖可得知在 e3 "教師與學生間自治與自省"與 e4 "教學目標"中,整體偏向同意至非常同意。

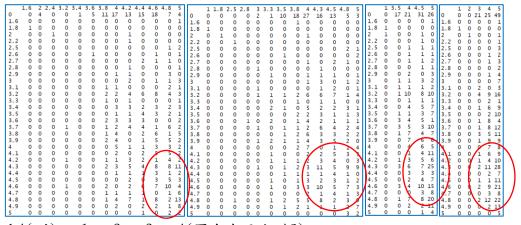
■ 自變數與應變數之二維列聯表

由於頻率的列聯表較不方便一次觀察,因此此部分只放上了以次數計 算的內容來一併說明,結果如下:



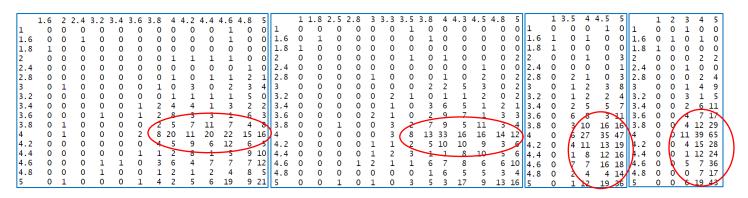
b3(列); e1、e2、e3、e4(圖由左至右-欄)

在 b3 與 e1、e2 的列聯表中,自變數與應變數的分數較為接近;而在 與 e3、e4 的表中,得知通常給予稍微較低的 b3 "網路社會資本中情感連 結型相關"分數時,也還是會給到較高的 e3 "教師與學生間自治與自省" 與 e4 "教學目標"分數。



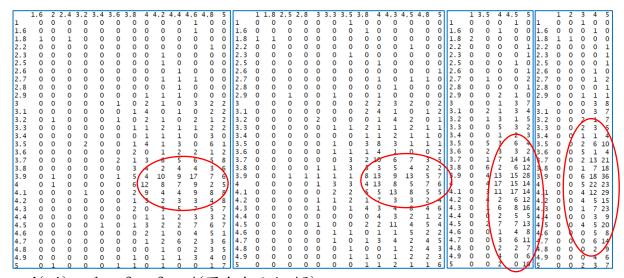
b4(列); e1、e2、e3、e4(圖由左至右-欄)

在自變數 b4 與各應變數的列聯表中,雖然資料數比較少,但可看出整體來說和前者相似,e1 與 e2 中自變數與應變數的分數較為接近;且在與 e3、e4 的表中,得知通常給予稍微較低的 b3 "網路社會資本中情感連結型相關"分數時,也還是會給到較高的 e3 "教師與學生間自治與自省"與 e4 "教學目標"分數。



c(列); e1、e2、e3、e4(圖由左至右-欄)

在自變數 c 與各應變數的列聯表中,整體來說與 el、e2 的自變數對應變數較無太大影響,大多"求助動機"為 4 分同意時,"教師與學生間之教學模式"與"校務參與"還是能獲得同意至非常同意的分數;而與 e3、e4 的表中,得知"求助動機"的分數,也對"教師與學生間自治與自省"與"教學目標"分數也無太大影響,皆偏向 4 至 5 分。



d(列); e1、e2、e3、e4(圖由左至右-欄)

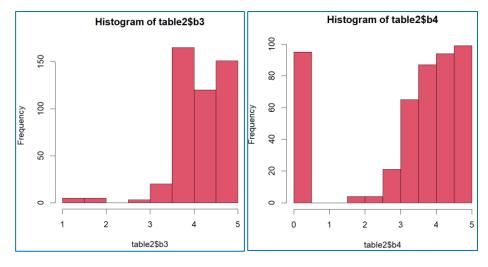
在自變數 d 與各應變數的列聯表中,即使給予"自我調整學習"4分左右的分數時,"教師與學生間之教學模式"與"校務參與"以及"教師與學生間自治與自省"與"教學目標"也會偏向4至5分。

整體來說這些自變數與應變數間其實較無明顯的線性關係。

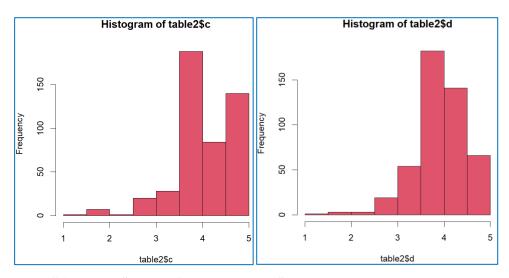
接下來則繪製視覺化圖形來更直觀的觀察各變數的狀態:

4. 視覺化圖形

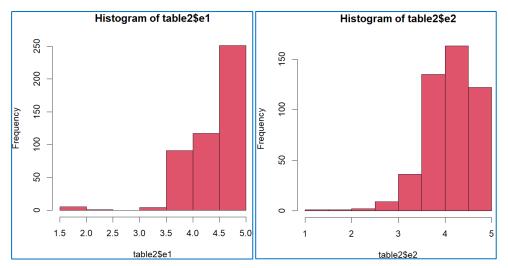
因為想以填答的分數高低來分析,因此將變數的 1~5 分取平均值後做 為連續型變數來觀察,以下為各變數之<mark>直方圖</mark>:



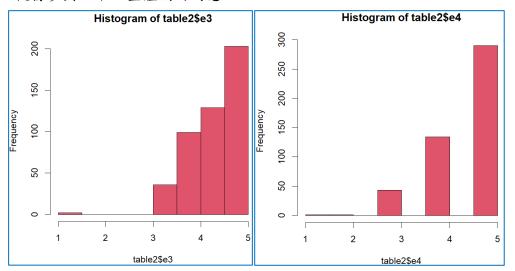
由上圖可得知在 b3 "網路社會資本中情感連結型相關"中,以 4 分最 多,次多的為 5 分;而 b4 "網路社會資本中任務導向型相關"中,分數越 高的比例越高,整體來說較偏同意至非常同意。



在 c "求助動機"與 d "自我調整學習"中,4 分的比例最多,而整體偏向同意。



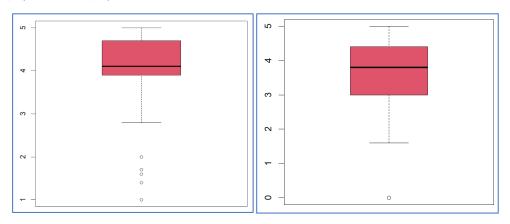
在 el "教師與學生間之教學模式"中,整體多為同意至非常同意;而在 e2 "校務參與"中,整體偏向同意。



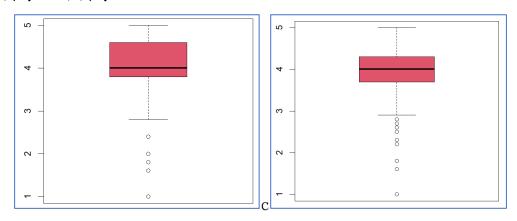
在 e3 "教師與學生間自治與自省"與 e4 "教學目標"中,整體偏向同意至非常同意,且多為非常同意。

箱形圖

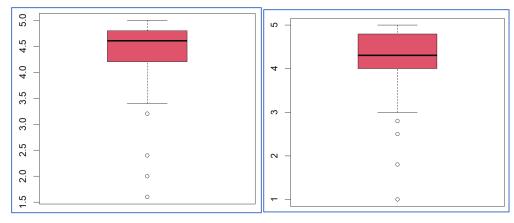
左圖為b3;右圖為b4。



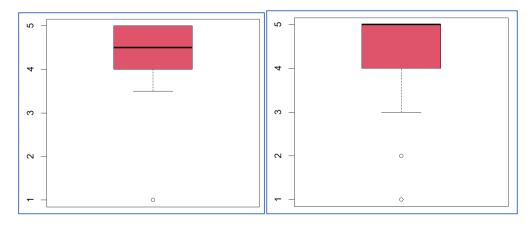
左圖為 c;右圖為 d。



左圖為 el;右圖為 e2。

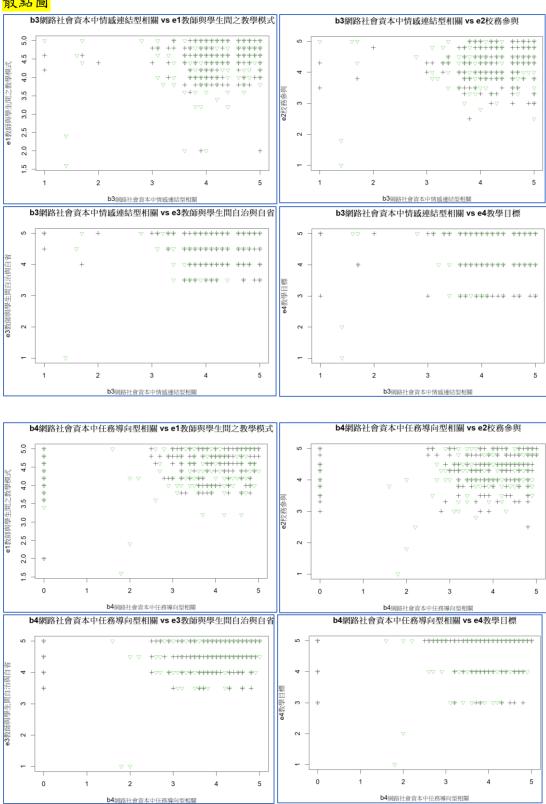


左圖為 e3;右圖為 e4。

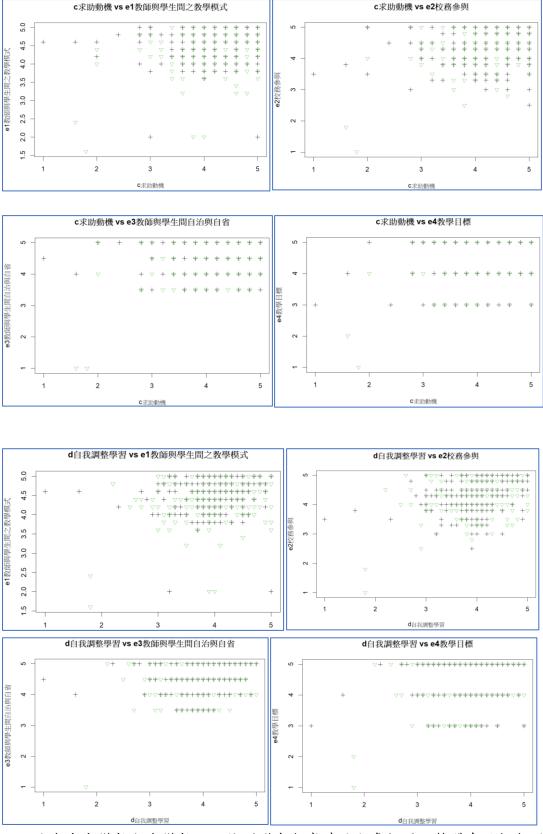


可知自變數的中位數大多在 4 分左右,應變數的中位數較高,大多在 4 至 5 分;且 b3、d 與 e1 的整體分數較集中;自變數落在範圍外的離群值較多,填答分數的範圍較廣。

散點圖



b4網路社會資本中任務導向型相關



此與在自變數和應變數之二維列聯表觀察時的大多相同,整體來說較無明 顯的線性關係。