

# 逢甲大學學生報告 ePaper

## 報告題名：

人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例

作者：李家淇、盧薪竹、范妍汝、張祐誠、張芸菁、陳宛謙、楊庭懿、  
賴皇鈞、鄭嘉程、柯雅潔、謝宇凡

系級：統計三甲、國貿全英四、應數四甲、財金四乙

學號：D0780997、D0571561、D0668835、D0679295、D0738736、  
D0739282、D0739323、D0739367、D0767675、D0780895、D0780906

開課老師：陳婉淑老師

課程名稱：迴歸分析

開課系所：統計系

開課學年： 109 學年度 第 一 學期

## 中文摘要

本研究旨在瞭解逢甲大學工科院大三與大四學生就讀研究所的意願及影響就讀研究所意願的潛在因素。在此研究中，本組設計調查問卷，於逢甲大學工科院大三與大四學生中隨機發放，按照科系人數及男女比例進行抽樣調查。最終將收集到的問卷結果，運用 SAS 進行多元迴歸分析及診斷分析，例如：殘差的獨立性與常態檢定；並運用 R、Python 與 Excel 將資料視覺化呈現，例如：散佈圖、堆疊長條圖、盒型圖、相關係數熱力圖等圖表，進而發展敘述統計分析。分析結果顯示，逢甲大學工科院大三與大四學生就讀研究所意願，對於疫情、研究所費用支出、研究所學術環境、實習機會與職業發展前景、缺乏軟硬實力認知、大學期間是否獲得書卷獎呈現正相關；對於增廣人脈、研究所往返住家距離、是否參加公務員補習呈現負相關。

**關鍵字：**就讀研究所的意願、相關性、多元迴歸分析



## Abstract

The purpose of this paper is to understand the junior and senior students' intention to study in graduate school and the factors influencing their intention to study in graduate school based in the Science and Engineering collage in Feng Chia University. In this study, our group designed questionnaires and sampled juniors and seniors from the Engineering college of Feng Chia University according to the number of departments and the ratio of male to female. Our group also used SAS stepwise selection method to conduct a multivariate regression analysis, ANOVA, and diagnostic analysis (independence, residual, and normality). In addition, using R, Python and Excel to make scatter plots, stacked bar charts, box charts, and correlation coefficient thermal charts for descriptive statistical analysis. The results of the research and analysis show that there is a positive correlation between intention to study in the graduate school and the COVID-19, the expenses of the graduate school, the academic environment of the graduate school, the internship opportunity and the career development, the lack of capability, and the times of awarding the best performance honor during the university. There was a negative correlation with the increased social network, the distance between graduate school and home, participation in civil service tutorials, or non-participation in tutorials.



**Keyword :** Intention to study in graduate school, Correlation, Multivariate linear analysis

## 目錄

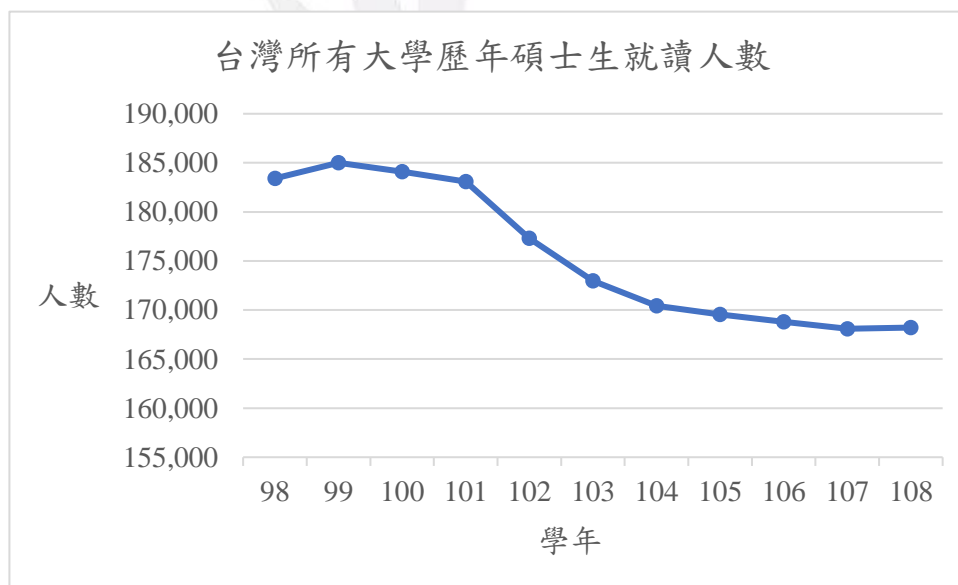
第一章、緒論.....	4
第二章、研究方法與模型.....	6
第一節、研究過程.....	6
第二節、研究方法.....	6
第三章、敘述統計.....	10
第一節、資料介紹.....	10
第二節、介紹變數.....	11
第三節、相關程度.....	14
第四節、敘述統計.....	18
第五節、獨立樣本 t 檢定 .....	21
第四章、迴歸分析.....	26
第一節、變數選取.....	26
(一)、多重共線性 .....	26
(二)、選取法 .....	27
(三)、有界變數轉換(Bounded Variable Transformation).....	27
第二節、配適模型.....	29
第三節、最終模型.....	32
(一)、檢定殘差的平均數為零 .....	33
(二)、檢定殘差之間是否獨立 $Cov\epsilon_i, \epsilon_j = 0, \forall i \neq j$ .....	33
(三)、檢定殘差是否服從常態分布 .....	34
第四節、迴歸方程式.....	35
第五章、結論.....	36
參考資料.....	37

## 第一章、緒論

近幾年，隨著時代的發展與就業市場對學歷的要求提高，即將畢業的大學生，都開始有「是否要繼續就讀研究所」的疑問。多花一、兩年的時間讀研究所，真的會比較好嗎？會比較容易找到工作嗎？碩士畢業的薪水會因此比學士畢業的薪水高嗎？若僅是高於兩、三千的薪水，是不是直接就業比較有優勢？

現今，大學學歷變得極度普遍，而漸漸地碩士學歷開始成為一種基本學歷的門檻。是不是只要家人、老師與同儕的鼓勵，就必須就讀研究所？這些原因都是直接或間接地導致大學生就讀研究所的因素。曾經就讀研究所是極為困難的事情，大學生為了提高自身能力與價值選擇就讀研究所。但如今卻是，就讀研究所不一定會比較好，大學畢業並進入社會職場，得到的經驗也比研究所多，在這樣的優勢，想要就讀研究所的人數就開始下降了。

根據圖(一)，台灣所有大學歷年就讀研究所的人數確實是呈現下降的趨勢，其原因為招生氾濫造成素質下滑這項因素。可以從圖(一)；教育部-統計處發現 99 學年度的研究所就讀人數創下史上高峰，超過 18 萬 5000 人。不管是否需要研究所學歷，只要親人的鼓勵或同儕的報名，就會報考研究所。但結果是，愈來愈多人懷疑研究所的價值，尤其當業界發現碩士生程度似乎不如預期時，大學生就讀研究所的意願開始降低，導致 100 學年後就讀研究所人數歷年降低。



圖(一) 台灣所有大學歷年碩士生就讀人數

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

為了證實圖(一)，我們藉由發放問卷去尋問逢甲大學工程與科學學院的大三和大四學生關於他們對未來是否就讀研究所的意願，其中工科院大三與大四的所有學生，總人數為 2164 人，抽樣人數為 165 人；調查系所分別為：機械與電腦輔助工程學系、纖維與複合材料學系、工業工程與系統管理學系、化學工程學系、應用數學系、航太與系統工程學系、環境工程與科學學系、材料科學與工程學系、光電科學與工程學系、精密系統設計學士學位學程，共 10 種不同科系做調查，並依照各科系總人數之比例做分配發放問卷，其中占母體人數 (2164) 7.62%。

近期，我們大三生開始面臨到，是否要就讀研究所的問題。在我們身邊的同儕慢慢開始確定目標，但也不僅是我們系上的大三生將要面對的問題。是全校、全台灣的大三生都開始正視的問題。而今年大力推廣五年一貫制，讓想讀研究所的同學，更有意願就讀研究所。但並非所有人都因此而就讀研究所。我們想要瞭解，實際的意願程度，透過發放問卷給逢甲大學工科院三四年級的學生，進行調查與研究，他們是否有意願就讀研究所、探討大三大四對於就讀研究所的看法以進而統計分析出大三大四生會因何種因素提高就讀研究所的意願。從問卷中可以瞭解逢甲大學工科院三四年級的學生的真實想法與情況，由此讓學校以及相關的教育部門可以對此做出更多的相關對策與方案。



## 第二章、研究方法與模型

### 第一節、研究過程

使用 SAS、R 與 Python 三種軟體進行敘述統計與迴歸分析。

1. SAS：進行選取變數迴歸分析與診斷分析(獨立性、殘差與常態)

2. R：製作相關係數矩陣熱力圖與盒型圖

- 相關係數矩陣熱力圖：透過計算各個變量得到相關係數，將獲得的數值透過矩陣熱力圖的方式呈現。熱力圖可以透過顏色變化反應矩陣中數據的訊息，顏色深的代表相關性越高；反之，顏色淺的代表相關性越低。熱力圖以直觀的方式把相關程度以顏色深淺標示出來。
- 盒型圖：是一種用來顯示數據分散情況的統計圖，在盒型圖中我們可以容易的發現全部數據的最大值、最小值、中位數、第一四分位數和第三四分位數，並可以了解數據中有無離群值和數據的偏態係數。

3. Python：製作長條圖、堆疊長條圖以及散佈圖。

- 長條圖：是一種以長方形的長度為變量的統計圖，長條圖通常可以用來比較兩個或兩個以上不同的條件或時間，只有一個變量，適用於較小的數據分析。
- 堆疊長條圖：是一種長條圖的變化形式，堆疊長條圖適合使用在複選題的數據表示，可以更加明確的比較一個數據中不同類型的差異，且可以比較它們總數之間的差異。
- 散佈圖：在迴歸分析中樣本點以直角坐標系為基準的平面上的一種分佈圖。散佈圖表示應變量將隨自變量而變化，所呈現的趨勢。我們可以利用散佈圖選擇相對合適的函數，對我們所擁有的數據點進行線性分布。

### 第二節、研究方法

1. 向前選取法(Forward Selection)：

一開始沒有自變數，每次選取都只加入一個自變數，自變數一旦加入到模型中，就不再剔除。將顯著值小於 0.1 的自變數加入模型，直到所有顯著的自變數都在模型中，便停止選取。

2. 向後消去法(Backward Elimination)：

一開始先把所有的自變數放進迴歸模型中，並且每一次只刪除一個自變數，一旦將變數刪除，就不可再放回迴歸模型中。把顯著值大於 0.10 的變數從模型剔除，直到所有不顯著的自變數都不在模型裡後便停止。



3. 逐步迴歸法(Stepwise Regression)：

一開始沒有自變數，每一個步驟中，會加入顯著值小於 1% 的自變數。若選入的自變數，顯著值大於 1.5%，將會剔除此自變數。直到沒有任何自變數被加入或剔除，則停止選取。

4. 變異數膨脹因子(Variance Inflation Factor；VIF)：

透過觀察解釋變數的數值來計算它們的多重共線性程度的數值，其公式如下：

$$VIF = \frac{1}{1-R_j^2}, j = 1, \dots, p-1,$$

( $R_j$  為解釋變數  $x_j$  對其他的解釋變數做迴歸分析時的相關係數)

一般來說，判斷準則為  $VIF$  介於 0 到 10 之間表示解釋變數沒有多重共線性； $VIF$  大於 10 表示解釋變數有多重共線性。

5. 相關係數(Correlation Coefficient)：

相關係數是由統計學家卡爾·皮爾森 Karl Pearson 提出，用來研究各個變數之間的線性關係，其公式如下：

$$r(x, y) = \frac{Cov(x, y)}{\sqrt{Var(x)Var(y)}},$$

其中， $x$  為自變數， $Var(x)$  為  $x$  的變異數， $y$  為應變數， $Var(y)$  為  $y$  的變異數， $Cov(x, y)$  為共變異數。

如果  $r(x, y)$  是正值說明它們正相關，反之是負值則有負相關， $|r(x, y)|$  越大說明它們相關性越大，如果  $|r(x, y)|$  越接近 0 則相關性越小。

6. 殘差常態分布假設：

我們使用四種常態假設檢定，分別是：Shapiro-Wilk test、Kolmogorov-Smirnov test、Cramer-von Mises test 以及 Anderson-Darling test，其假設如下：

$H_0$ ：The data follows normal distribution，

$H_a$ ：The data do not follow normal distribution。

● Shapiro-Wilk test：

用來檢定當樣本數在 50 個以下是否為常態分布，其公式如下：

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

其中， $x_i$  為樣本值， $x_{(i)}$  為第  $i$  階層樣本數， $\bar{x}$  為  $x_i$  樣本值的平均數， $a_i$  為第  $i$  階層的係數。

若檢定之  $p$ -value 小於 0.05，表示拒絕虛無假設，資料不符合常態分布；反之  $p$ -value 大於 0.05，則表示不拒絕虛無假設，資料符合常態分布。

● Kolmogorov-Smirnov test：

用來檢定當樣本數在 50 個以上是否為常態分布，其公式如下：

$$K = \sup_{t \in [0,1]} |B(t)|,$$



其中， $B(t)$ 為布朗橋， $K$ 的累積分布函數為：

$$P_r(K \leq x) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 x^2}。$$

若檢定之 $p-value$ 小於 0.05，表示拒絕虛無假設，資料不符合常態分布；反之 $p-value$ 大於 0.05，則表示不拒絕虛無假設，資料符合常態分布。

● Anderson-Darling test：

用來檢定是否為常態分布，其公式如下：

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln F(x_i) + \ln(1 - F(x_{n-i+1}))]，$$

其中， $x_i$ 為樣本值， $n$ 為樣本數， $F(x_i)$ 為常態分布的累積分布函數。

若檢定之 $p-value$ 小於 0.05，表示拒絕虛無假設，資料不符合常態分布；反之 $p-value$ 大於 0.05，則表示不拒絕虛無假設，資料符合常態分布。

● Cramer-von Mises test：

用來檢定是否為常態分布，其公式如下：

$$T = \frac{NM}{N+M} \omega^2，$$

其中， $N$ 為第一個樣本數， $M$ 為第二個樣本數， $\omega$ 為估計經驗分布函數和真實經驗分布函數的差距。

若檢定之 $p-value$ 小於 0.05，表示拒絕虛無假設，資料不符合常態分布；反之 $p-value$ 大於 0.05，則表示不拒絕虛無假設，資料符合常態分布。

7. 殘差自我相關檢定(Durbin-Watson)：

主要檢測迴歸分析中殘差項是否存在自我相關，其公式如下：

$$D = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}，$$

Durbin-Watson 分別檢定自我相關是否在顯著水準為正與負，其檢定如下：

$$H_0 : \rho_1 = 0，$$

$$H_a : \rho_1 > 0，$$

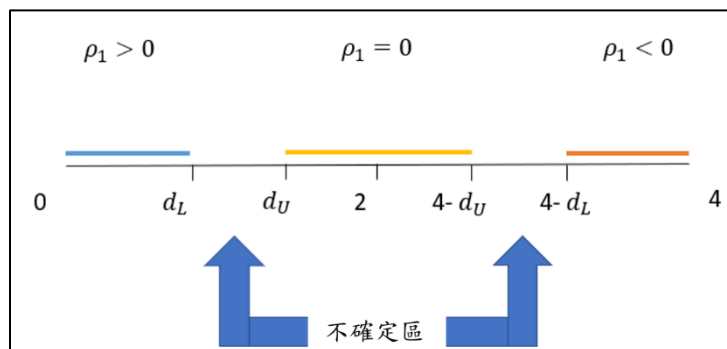
若正的自我相關檢定 $D$ 小於 $d_{L,a}$ ，表示拒絕虛無假設；若正的自我相關檢定 $D$ 大於 $d_{U,a}$ 時，表示不拒絕虛無假設；若負的自我相關檢定 $D$ 介於 $d_{L,a}$ 與 $d_{U,a}$ ，則無法定論。

$$H_0 : \rho_1 = 0$$

$$H_a : \rho_1 < 0$$

若負的自我相關檢定 $(4-D)$ 小於 $d_{L,a}$ ，表示拒絕虛無假設；若負的自我相關檢定 $(4-D)$ 大於 $d_{U,a}$ 時，表示不拒絕虛無假設；若負的自我相關檢定 $(4-D)$ 介於 $d_{L,a}$ 與 $d_{U,a}$ ，則無法定論。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)



圖(二)殘差的自我相關檢定範圍

由圖(二)可知，當檢定結果接近0，有負自我相關；當檢定結果接近4，有正自我相關；當檢定結果接近2，則為無自我相關。

#### 8. 獨立樣本 t 檢定(Independent-Sample t test)：

主要比較兩組樣本的平均值之間是否存在顯著差異，其公式如下：

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma x_1^2 + \sigma x_2^2 - 2\gamma \sigma x_1 \sigma x_2}{n-1}}},$$

檢定是否存在顯著差異的提前是：應變數必須是連續的且從母體中隨機抽樣的樣本，而母體必須服從常態分布、樣本必須是獨立事件，兩組樣本的變異數須為常態分布。

$$H_0 : \mu_1 = \mu_2,$$

$$H_a : \mu_1 \neq \mu_1,$$

首先必須檢驗變異數是否相等，若變異數不相等  $p$ -value 小於顯著水準，將確認方法 Satterthwaite 是否拒絕虛無假設；若變異數相等  $p$ -value 大於顯著水準，將確認方法 Pooled 是否拒絕虛無假設。

## 第三章、敘述統計

### 第一節、資料介紹

此研究的母體為逢甲大學工程與科學學院大三與大四的所有學生，母體總人數為 2164 人，隨機抽取 165 人為樣本，並依照工科院各科系總人數、年級人數、男女人數之比例，將人數平均分配至大三與大四的學生，而男女人數是依照 107 學年總人數之比例進行分配。表(一)為依照工科院母體比例之問卷人數分配表。實際發放問卷的人數，年級與男女比例有些許差異。例如：機電系大三女生之問卷有 3 份、機電系大四女生問卷為 0。

表(一)依照工科院母體比例之問卷人數分配表

科系	總人數	百分比	大三人數	大四人數	男生人數 (大三+大四)	女生人數 (大三+大四)
機械與電腦輔助工程學系	25	1.16%	14	11	11+11	3+0
纖維與複合材料學系	16	0.74%	8	8	6+6	2+2
工業工程與系統管理學系	16	0.74%	10	6	6+4	4+2
化學工程學系	18	0.83%	9	9	6+6	3+3
應用數學系	14	0.65%	7	7	6+6	1+1
航太與系統工程學系	19	0.88%	10	9	9+8	1+1
環境工程與科學學系	16	0.74%	7	9	5+6	2+3
材料科學與工程學系	17	0.79%	10	7	6+7	4+0
光電科學與工程學系	16	0.74%	8	8	7+7	1+1
精密系統設計學士學位學程	8	0.37%	4	4	3+3	1+1
總計	165	7.62%	87	78	129	36

## 第二節、介紹變數

- y：我願意花 1~2 年的時間就讀研究所，投資自己的未來。(意願程度：1~10)
- x1：受 COVID-19 疫情的影響，新鮮人求職環境欠佳，我認為就讀研究所是較好的選擇。(意願程度：1~10)
- x2：我清楚了解報考研究所的管道、時間和考試科目。(意願程度：1~10)
- x3：我認為未來嚮往的目標職場對於研究所學歷有需求。(意願程度：1~10)
- x4：我認為就讀研究所所需的費用支出（例如：學費、生活費）是有價值的。(意願程度：1~10)
- x5：我認為研究所的學術環境(例如:圖書館、實驗室、研討室)是重要的。(意願程度：1~10)
- x6：我認為讀研究所能增廣人脈，對未來工作上會有幫助。(意願程度：1~10)
- x7：對於我就讀的科系，比起更高學歷，提早進入職場累積實務經驗對未來更有幫助。(意願程度：1~10)
- x8：我認為自身就讀研究所的意願或動機會受到他人影響(如：家人、師長或同儕)。(意願程度：1~10)
- x9：相較於大學部，我認為研究所能提供更好的實習機會與職業發展前景。(意願程度：1~10)
- x10：我清楚了解逢甲大學研究生獎助學金辦法。凡參加逢甲碩士班甄試入學及碩士班考試入學之正取生，前七學期學業成績總平均名次為該系、學位學程前百分之六以內者，給予第一學年全額學雜費助學金；其於本校前七學期學業成績總平均名次為該系、學位學程前百分之六至百分之二十以內者，給予第一學年半額學雜費助學金。(意願程度：1~10)
- x11：我認為逢甲大學五年一貫(建築系六年一貫)的預研究生制度會提高我就讀研究所的意願。(意願程度：1~10)
- x12：若逢甲大學提供逢甲研究所升學管道說明會我有意願參加。(意願程度：1~10)
- x13：與其就讀國立大學研究所兩年，我傾向花一年的時間以預研究生的身分就讀逢甲研究所。(意願程度：1~10)
- x14：我認同逢甲大學研究所的聲譽與排名。(意願程度：1~10)
- x15：我認同逢甲大學的師資與研究重點。(意願程度：1~10)
- x16：我認為就讀研究所會讓他人對於自身的觀感提升(如：尊重程度、社會地位)。(意願程度：1~10)
- x17：我認為就讀國立大學碩士班所取得的學歷較具競爭力。(意願程度：1~10)
- x18：尚未知道出社會所需之硬實力與軟實力，因此選擇就讀碩士班加強技能。(意願程度：1~10)
- x19：在學期間，我的班排名次平均落在：(1：前 1% 到 20%，2：前 21% 到 40%，3：前 41% 到 60%，4：前 61% 到 80%，5：前 81% 到 100%)

(人生轉捩點-大學畢業後要繼續就讀研究所嗎?以逢甲工科院為例)

- x20：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：國/私立。  
(1：勾選 0：未勾選)
- x21：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：住家距離。  
(1：勾選 0：未勾選)
- x22：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：當地物價。  
(1：勾選 0：未勾選)
- x23：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：產學合作/實習。  
(1：勾選 0：未勾選)
- x24：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：課程配置。  
(1：勾選 0：未勾選)
- x25：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：獎學金。  
(1：勾選 0：未勾選)
- x26：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：學校資源。  
(1：勾選 0：未勾選)
- x27：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：學校師資。  
(1：勾選 0：未勾選)
- x28：若我要就讀研究所，我將優先考慮的條件(最多複選3個)：業界知名度。  
(1：勾選 0：未勾選)
- x29：大學期間，是否參加以下校外學習(可複選)：研究所考試科目補習。  
(1：勾選 0：未勾選)
- x30：大學期間，是否參加以下校外學習(可複選)：外語文補習。  
(1：勾選 0：未勾選)
- x31：大學期間，是否參加以下校外學習(可複選)：本科系專業證照。  
(1：勾選 0：未勾選)
- x32：大學期間，是否參加以下校外學習(可複選)：公務員高、普、特考補習。  
(1：勾選 0：未勾選)
- x33：大學期間，是否參加以下校外學習(可複選)：無。(1：勾選 0：未勾選)
- x34：大學期間，我曾獲得書卷獎的次數：(0：零次，1：一次，2：兩次，3：三次，4：四次，5：五次，6：六次，7：七次，8：八次，9：九次以上)
- x35：我認為研究所畢業後的起薪(單位：新台幣)。(EX：35000)
- x36：受訪者就讀的科系。(27：機電系 28：織複系 29：工工系 30：化工系  
31：應數系 32：航太系 33：環科系 34：材料系 35：光電系 36：精密學程)
- x37：受訪者就讀的年級。(0：大三，1：大四)
- x38：受訪者生理的性別。(0：男，1：女)
- D19：在學期間，我的班排名次平均落在。(0：前1%到40%，1：前41%到100%)
- D34：大學期間，我曾獲得書卷獎的次數。(0：無，1：有)

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

D1 至 D9：

表(二)變數 36 的虛擬變數

科系	D1	D2	D3	D4	D5	D6	D7	D8	D9
機電系	1	0	0	0	0	0	0	0	0
纖複系	0	1	0	0	0	0	0	0	0
工工系	0	0	1	0	0	0	0	0	0
化工系	0	0	0	1	0	0	0	0	0
應數系	0	0	0	0	1	0	0	0	0
航太系	0	0	0	0	0	1	0	0	0
環科系	0	0	0	0	0	0	1	0	0
材料系	0	0	0	0	0	0	0	1	0
光電系	0	0	0	0	0	0	0	0	1
精密學程	0	0	0	0	0	0	0	0	0



### 第三節、相關程度

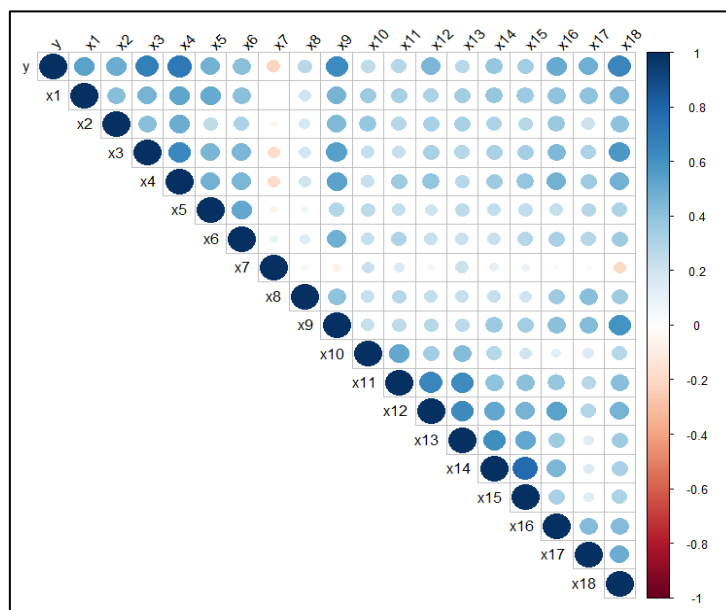
此章節的目的為探討 y 與變數 x1 到變數 x18 兩兩變數之間的相關性。透過 R 語言來建立表(三)相關係數矩陣與圖(三)相關係數矩陣熱力圖。主要以下三角的相關係數判斷第一部分 19 個變數的相關性。相關係數絕對值若介於 0.70 至 0.99，即代表兩變數具有高度相關性。從表(三)可以發現，y 和變數 x4 以及變數 x14 和變數 x15 的相關係數為 0.71 與 0.78，兩組變數之間具有高度相關性。相關係數絕對值若介於 0.49 至 0.69，即代表兩變數具有中度相關性。y 和變數 x3、y 和變數 x18 以及變數 x11 和變數 x12 的相關係數分別為 0.68、0.66 以及 0.66，三組變數之間具有中度相關。而無相關性的僅有一組變數 x1 和變數 x7，其相關係數為 0。出現負相關的結果，只發生在變數 x7，共有 9 組負相關。原因為變數 x7 是唯一的反向題，因此可以推論受訪者均認真思考並作答，才得到負相關結果。

表(三)相關係數矩陣

Column1	y	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18
y	1	0.54	0.5	0.68	0.71	0.47	0.41	-0.21	0.28	0.62	0.25	0.29	0.46	0.28	0.39	0.35	0.5	0.49	0.66
x1	0.54	1	0.43	0.46	0.52	0.5	0.42	0	0.21	0.46	0.35	0.34	0.32	0.35	0.38	0.36	0.41	0.4	0.46
x2	0.5	0.43	1	0.43	0.5	0.25	0.31	-0.06	0.17	0.45	0.38	0.29	0.33	0.32	0.31	0.28	0.37	0.22	0.4
x3	0.68	0.46	0.43	1	0.65	0.45	0.46	-0.19	0.19	0.55	0.23	0.24	0.33	0.29	0.33	0.33	0.45	0.32	0.57
x4	0.71	0.52	0.5	0.65	1	0.47	0.46	-0.19	0.2	0.54	0.23	0.36	0.4	0.28	0.36	0.4	0.47	0.36	0.48
x5	0.47	0.5	0.25	0.45	0.47	1	0.52	-0.08	0.08	0.29	0.27	0.24	0.21	0.25	0.24	0.25	0.24	0.28	0.3
x6	0.41	0.42	0.31	0.46	0.46	0.52	1	0.09	0.14	0.49	0.23	0.31	0.22	0.22	0.23	0.28	0.33	0.29	0.35
x7	-0.21	0	-0.06	-0.19	-0.19	-0.08	0.09	1	0.05	-0.09	0.21	0.15	0.06	0.22	0.11	0.09	-0.05	-0.04	-0.2
x8	0.28	0.21	0.17	0.19	0.2	0.08	0.14	0.05	1	0.4	0.24	0.28	0.24	0.23	0.24	0.2	0.36	0.42	0.35
x9	0.62	0.46	0.45	0.55	0.54	0.29	0.49	-0.09	0.4	1	0.23	0.26	0.28	0.26	0.36	0.34	0.42	0.43	0.6
x10	0.25	0.35	0.38	0.23	0.23	0.27	0.23	0.21	0.24	0.23	1	0.52	0.34	0.43	0.29	0.19	0.12	0.16	0.28
x11	0.29	0.34	0.29	0.24	0.36	0.24	0.31	0.15	0.28	0.26	0.52	1	0.66	0.63	0.41	0.42	0.38	0.28	0.42
x12	0.46	0.32	0.33	0.33	0.4	0.21	0.22	0.06	0.24	0.28	0.34	0.66	1	0.64	0.52	0.46	0.53	0.29	0.46
x13	0.28	0.35	0.32	0.29	0.28	0.25	0.22	0.22	0.23	0.26	0.43	0.63	0.64	1	0.62	0.52	0.35	0.14	0.35
x14	0.39	0.38	0.31	0.33	0.36	0.24	0.23	0.11	0.24	0.36	0.29	0.41	0.52	0.62	1	0.78	0.45	0.15	0.33
x15	0.35	0.36	0.28	0.33	0.4	0.25	0.28	0.09	0.2	0.34	0.19	0.42	0.46	0.52	0.78	1	0.32	0.14	0.32
x16	0.5	0.41	0.37	0.45	0.47	0.24	0.33	-0.05	0.36	0.42	0.12	0.38	0.53	0.35	0.45	0.32	1	0.44	0.44
x17	0.49	0.4	0.22	0.32	0.36	0.28	0.29	-0.04	0.42	0.43	0.16	0.28	0.29	0.14	0.15	0.14	0.44	1	0.49
x18	0.66	0.46	0.4	0.57	0.48	0.3	0.35	-0.2	0.35	0.6	0.28	0.42	0.46	0.35	0.33	0.32	0.44	0.49	1



(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)



圖(三)相關係數矩陣熱力圖

表(四)相關性程度

相關性	變數
高度正相關： $r > 0.7$	y 與 x4 、 x14 與 x15
中度正相關： $r > 0.65$	y 與 x3 、 y 與 x18 、 x11 與 x12
無相關： $r = 0$	x1 與 x7
負相關： $r < 0$	x7 與 y、x2、x3、x4、x5、x9、x16、x17、x18

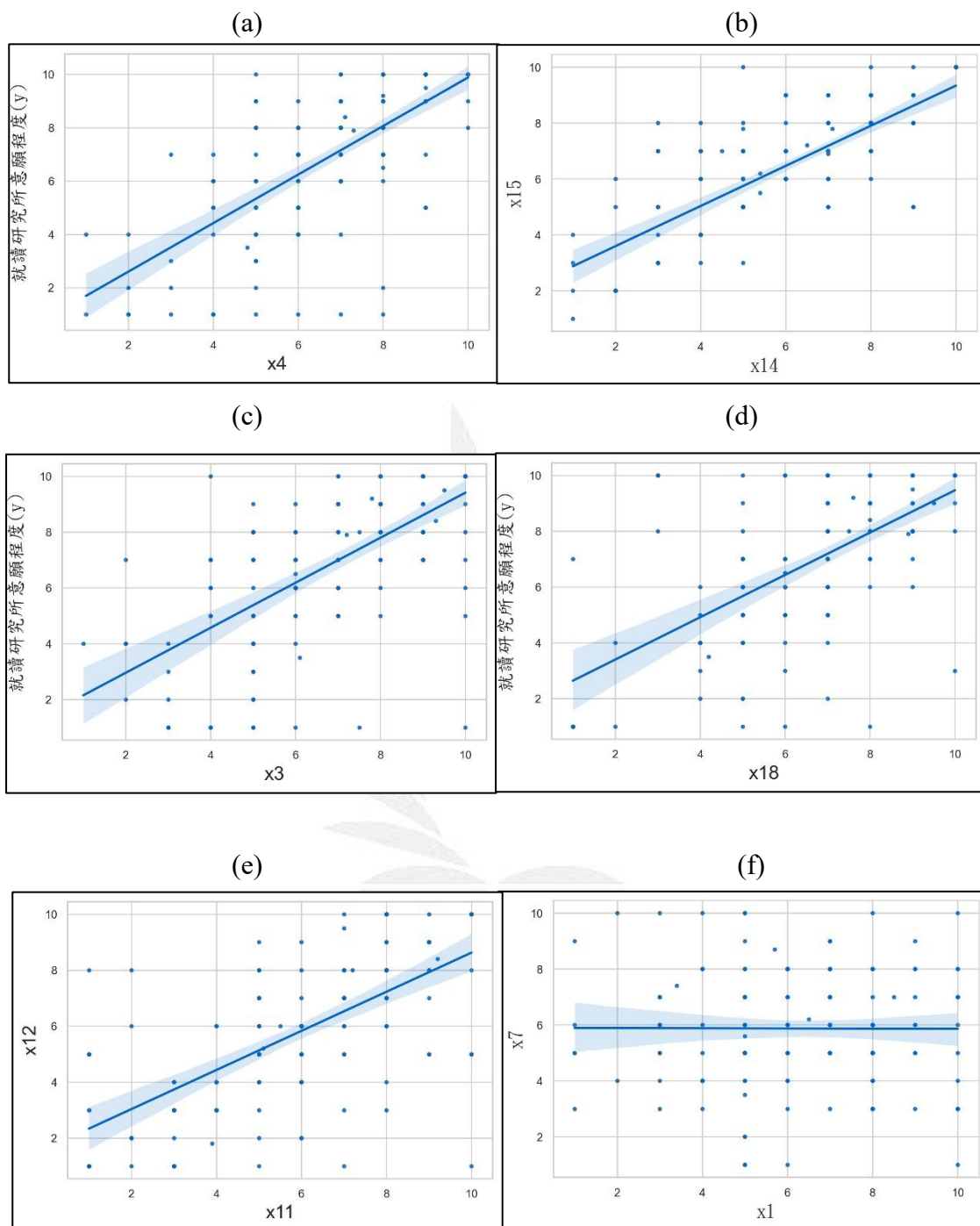
從表(四)分別探討變數之間具有高度正相關、中度正相關、無相關以及負相關。具有高度正相關的為 y 與變數 x4 以及變數 x14 與變數 x15。其中，y 與變數 x4 之間的相關係數為 0.71 具有高度正相關，由此我們可以推論，若受訪者願意就讀研究所，將就讀研究所需要的支出，例如：學雜費、生活費，視為合理或有價值。即使需要付出兩年就讀研究所的花費，學生也願意就讀研究所，從而學到專業學術能力及獨立自主的判斷力，這些能力在面對人生、公共議題或專業領域的工作時是有幫助的。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

圖(四) 相關係數散佈圖

(a) y 與 x4 (b) x14 與 x15 (c) y 與 x3

(d) y 與 x18 (e) x11 與 x12 (f) x1 與 x7



透過圖(四 a)看出 y 與 x4 具有高度正相關，我們可以推論，若受訪者願意就讀研究所，會將就讀研究所需要的支出，例如：學雜費、生活費，視為合理或有價值性的。透過圖(四 b)看出 x14 與 x15 具有高度正相關，我們可以推論，認同逢甲大學研究所聲譽與排名的學生，同樣也會認同逢甲大學的師資與研究重點。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

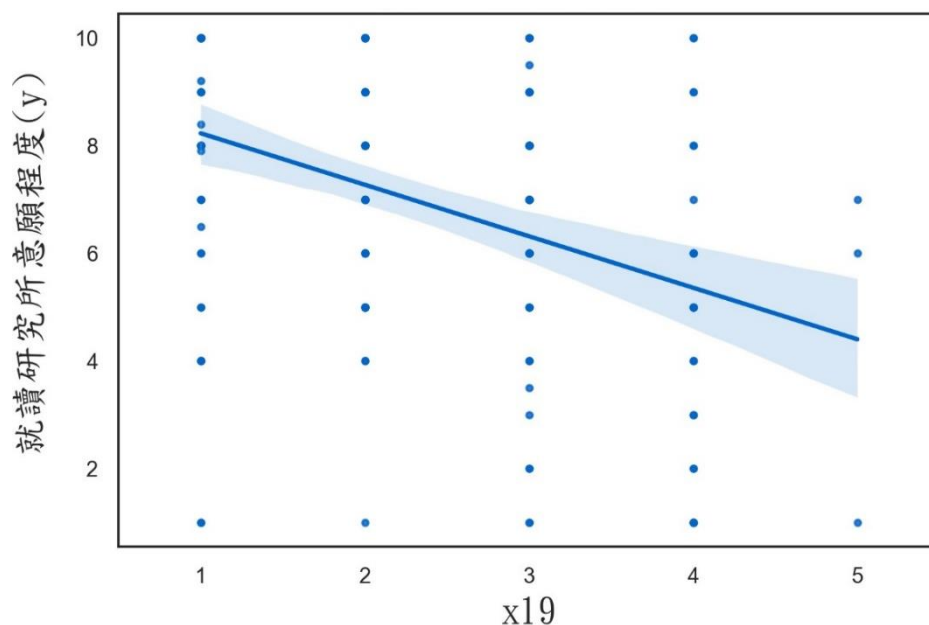
透過圖(四 c)看出 y 與 x3 具有中度正相關性。由此我們可以推論，多數受訪者認為就讀研究所在未來目標職場是有幫助的以及受訪者已查詢或接收相關資訊，因而提升學生就讀研究所的意願。透過圖(四 d)看出 y 與 x18 具有中度正相關。由此我們可以推論，學生追求更高的學歷和掌握更多的專業知識最主要目的無非是能夠獲得一份更好的工作。透過圖(四 e)看出 x11 與 x12 具有中度正相關。由此我們可以推論，越認同逢甲大學實行五年一貫（建築系六年一貫）的預研究生制度的學生，當學校提供逢甲研究所的升學管道說明會時，參加的意願會更加強烈。

透過圖(四 f)看出 x1 與 x7 不具相關性。由此我們可以推論，COVID-19 疫情並不會影響學生是否想要提早進入職場或選擇就讀研究所。



## 第四節、敘述統計

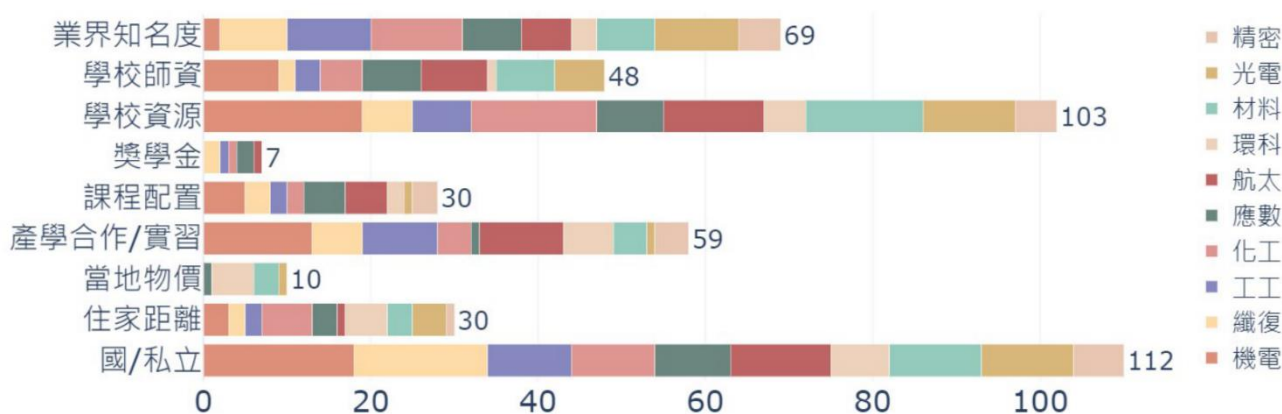
我們蒐集的問卷來源範圍遍佈大三與大四的工科院學生，包含十個不同科系。我們想透過敘述統計中的散佈圖、堆疊長條圖與群組盒型圖分析，知道不同年級和不同科系對於問題是否會有不同的看法。



圖(五) y 與變數 x19 的散佈圖

我們想要探究是否成績越好的受訪者越想要繼續就讀研究所，因此我們想要透過散佈圖探討這兩個變量之間的關係。圖(五)表示，變數 x19 與 y 呈現負相關，成績越好的受訪者就讀研究所的意願就越大。

### 第22題：若我要就讀研究所，我將先優先考慮的條件

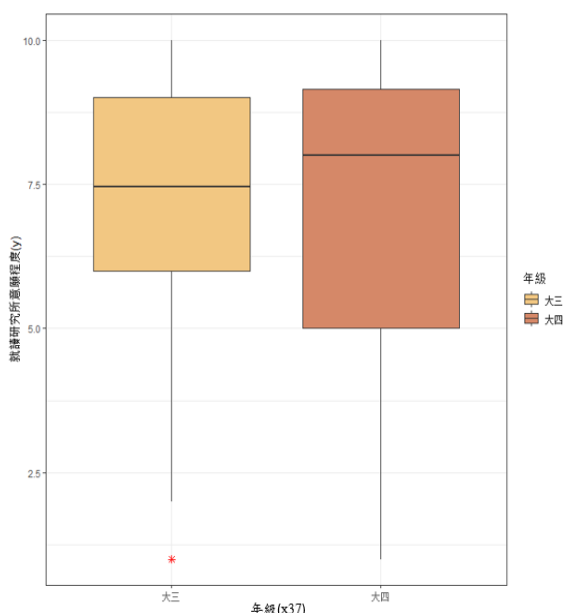


圖(六)各科系優先考慮研究所的條件

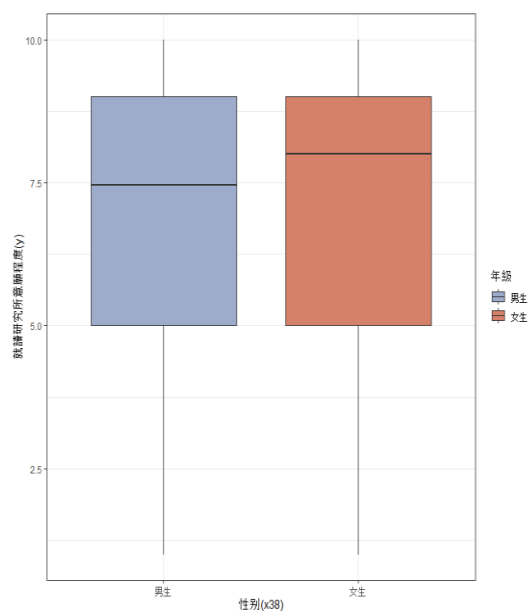
由圖(六)可知，從整體而言，工科院的學生在選擇就讀研究所優先考慮的前三個條件是國私立、學校資源以及業界知名度。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

化工系與材料系的學生較優先考慮學校資源與國(私)立，其中更優先考慮學校資源，代表學生比起就讀國立學校，更加注重學校資源是否足夠，例如：做實驗的器具或設備；航太系的學生較優先考慮學校資源、國(私)立與產學合作，代表航太系的學生非常注重學校資源和產學合作(實習)，例如：畫模型的程式、漢翔或中研院實習機會。

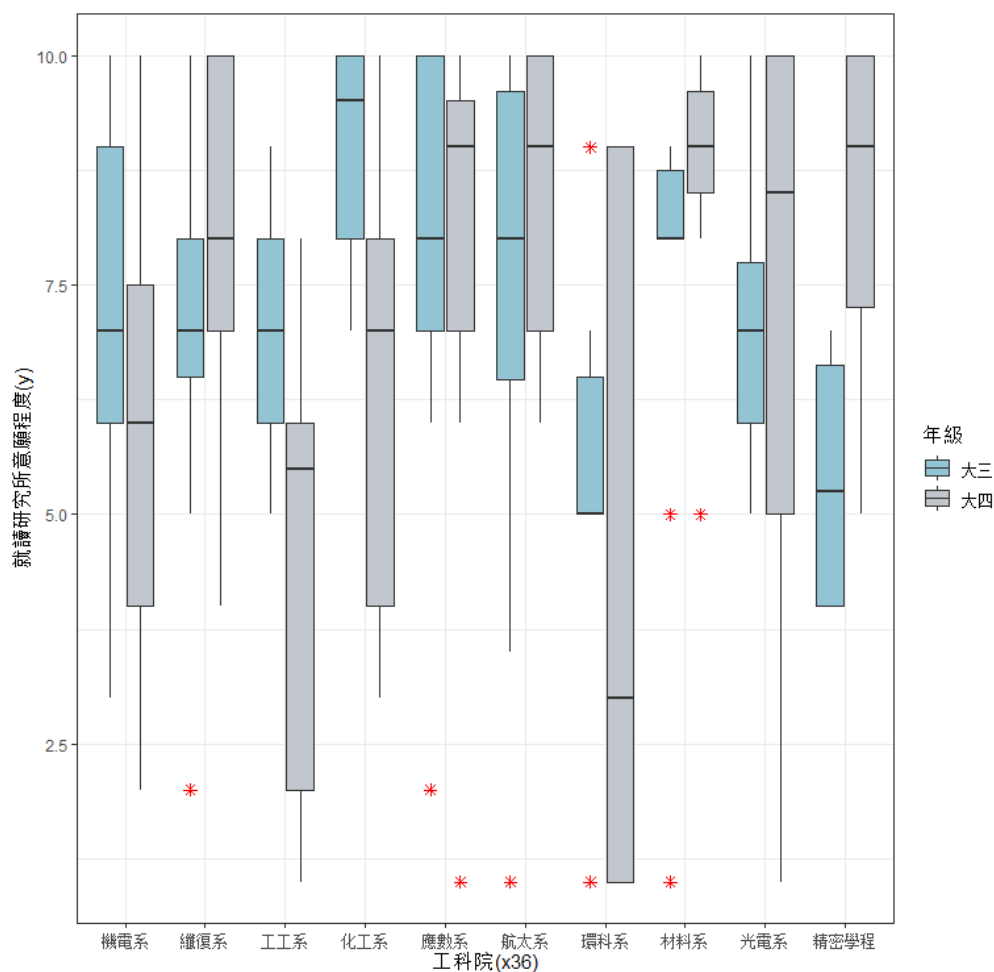


圖(七)年級與就讀研究所意願程度的盒型圖



圖(八)性別與就讀研究所意願程度的盒型圖

圖(七)與圖(八)分別是年級與性別對於就讀研究所的意願程度，我們可以發現，從年級觀察大三意願程度的全距比大四意願程度的全距小，可以推測：由於發放問卷時間已經是多數大四學生已知道研究所放榜結果，因此要就讀研究所的人數已有一部分確定就讀或是選擇不就讀而就業。從性別觀察女生意願程度的中位數高於男性，雖然女生的樣本數比較少，但從圖(八)可以發現，男女生對於就讀研究所的意願狀況是相似的。是否要就讀研究所的想法是類似的。



圖(九)各系就讀研究所的意願程度

透過圖(九)的盒型圖，我們可以發現，各系的大三與大四學生對於就讀研究所的意願情況。首先多數的意願程度中位數都高於 5，代表工科院大三與大四的學生對於就讀研究所的意願程度偏高。其次，從圖(九)可以發現材料系以及應數系的大三與大四學生，就讀研究所的意願程度平均比其他系所高。尤其是材料系，除了三個離群值以外的觀察值，全部都高於 7.5。

材料系就讀研究所的意願程度的全距最小，可以發現材料系的學生對於就讀研究所的意願程度皆大於 7.5。反之，環科系的學生對於就讀研究所的意願非常低，大三與大四就讀研究所的意願程度中位數都小於各系所的中位數且大四學生對於就讀研究所的意願程度的全距非常大，推論環科系大四學生對於自己未來的規劃比大三明確且兩極化。

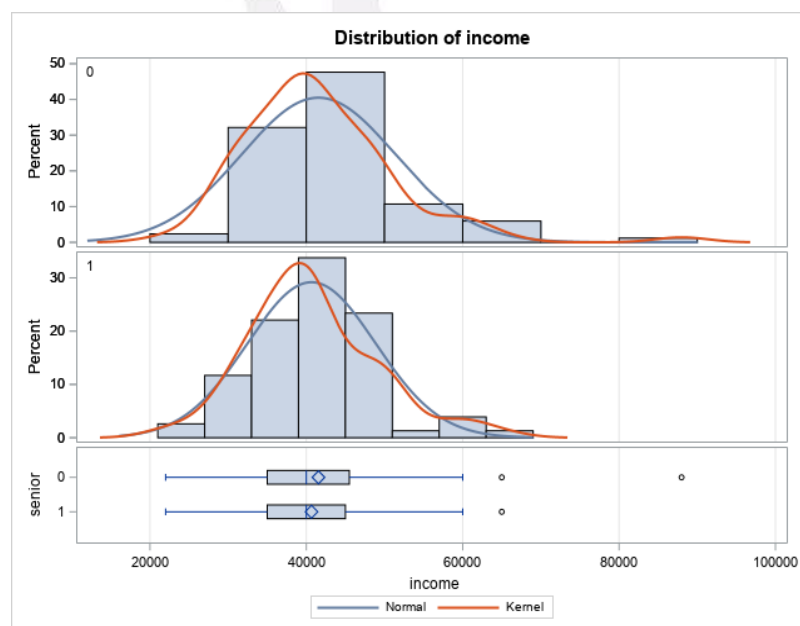
最後我們發現各系的大四就讀研究所的意願程度的全距皆大於該系大三的全距，這說明了大四的學生基本上已經決定自己是否要就讀研究所，才会有這麼兩極化的結果。

## 第五節、獨立樣本 t 檢定

從上述的敘述統計可以得知，各科系的男女、年級對於就讀研究所的意願以及對於研究所的優先選擇條件。每個科系未來的起薪都有所不同，不同年級或性別對於畢業後的起薪認知或許將有所不同；每個科系中成績的分布，分別以在班上名次的前後 40%與是否獲得書卷獎，或許就讀研究所的意願程度會有所不同。因此將以獨立樣本 t 檢定，觀察母體平均是否相等。

表(五)年級與起薪的獨立 t 檢定

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	159	0.62	0.5358
Satterthwaite	Unequal	157.58	0.63	0.5326
Method	Num DF	Den DF	F Value	Pr > F
Folded F	83	76	1.44	0.1061



圖(十)年級與起薪的獨立 t 檢定

兩組樣本分別是所有工科院大三學生希望的研究所畢業後的起薪和大四學生希望的研究所畢業後的起薪。我們想要探究這兩組樣本有無顯著差異，所以對這兩個樣本做獨立樣本 t 檢定。

$H_0$ ：大三和大四的學生對於他們的畢業起薪沒有顯著差異

$H_a$ ：大三和大四的學生對於他們的畢業起薪有顯著差異

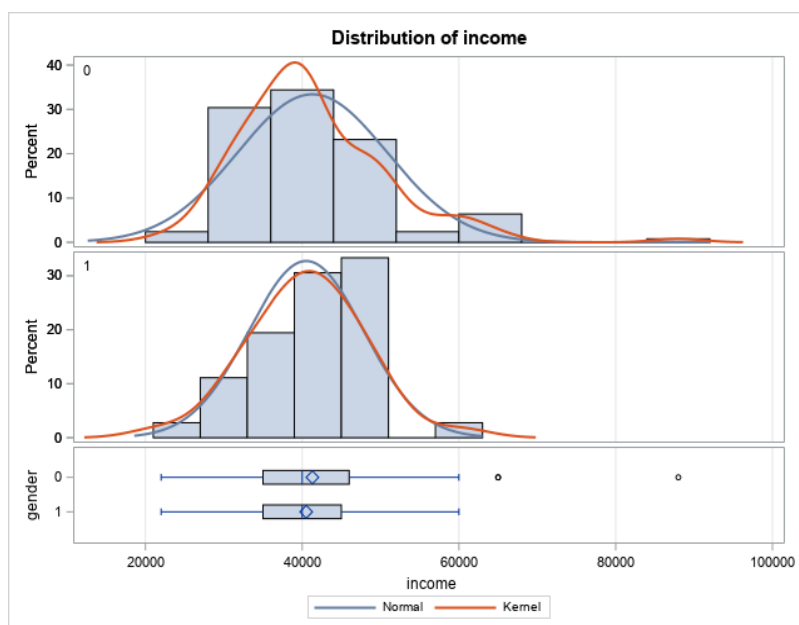


(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

經過分析表(五)的結果將如圖(十)所示。在 Equality of Variances 表中， $p - value$  為 0.1061，大於顯著水準  $\alpha$ ，所以要參考 Pooled 方法的  $p - value$ 。在圖(十)中，Pooled 的  $p - value$  為 0.5358，明顯大於顯著水準  $\alpha = 0.05$ 。則不拒絕虛無假設  $H_0$ ，說明大三大四的學生對於他們的畢業起薪無明顯差異。

表(六)性別與起薪的獨立 t 檢定

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	159	0.46	0.6471
Satterthwaite	Unequal	72.897	0.53	0.5972
Method	Num DF	Den DF	F Value	Pr > F
Folded F	124	35	1.71	0.0692



圖(十一)性別與起薪的獨立 t 檢定

兩組樣本分別是所有工科院中男生所希望的研究所畢業起薪和女生所希望的研究所畢業起薪。我們想要探究這兩組樣本有無顯著差異，所以對這兩個樣本做獨立樣本 t 檢定。

$H_0$ ：男生和女生對於他們的畢業起薪沒有顯著差異

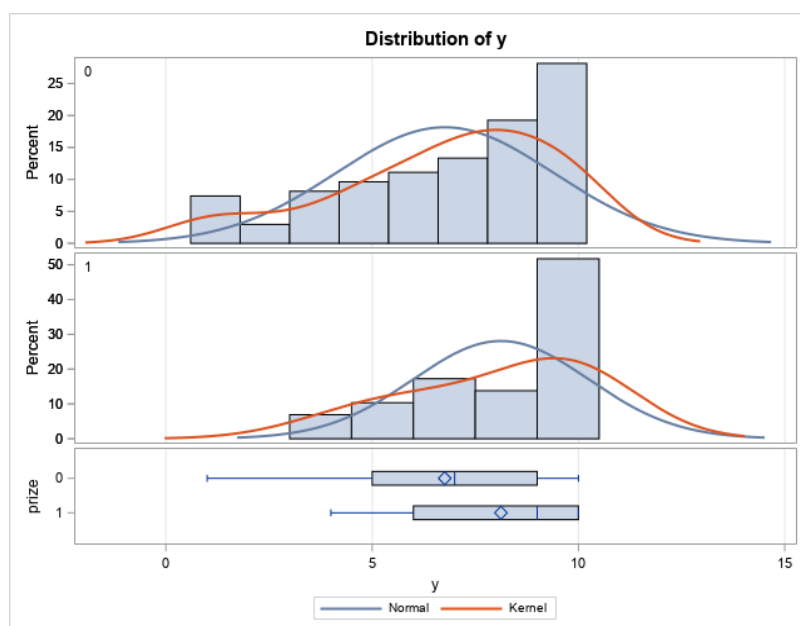
$H_a$ ：男生和女生對於他們的畢業起薪有顯著差異

經過分析表(六)的結果將如圖(十一)所示。在 Equality of Variances 表中， $p - value$  為 0.0692，明顯大於顯著水準  $\alpha = 0.05$ ，所以要參考 Pooled 方法的  $p - value$ 。在表(六)中，Pooled 的  $p - value$  為 0.6471，明顯大於顯著水準  $\alpha = 0.05$ 。則不拒絕虛無假設  $H_0$ ，說明男生和女生對於他們的畢業起薪無明顯差異。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

表(七)獲得書卷獎與就讀研究所意願的獨立 t 檢定

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	162	-2.60	0.0101
Satterthwaite	Unequal	48.354	-2.98	0.0045
Method	Num DF	Den DF	F Value	Pr > F
Folded F	134	28	1.53	0.1893



圖(十二)獲得書卷獎與就讀研究所意願的獨立 t 檢定

兩組樣本分別是所有工科院中得過書卷獎的受訪者想要念研究所的意願程度和沒有得過書卷獎的受訪者想要念研究所的意願程度。我們想要探究這兩組樣本有無顯著差異，所以對這兩個樣本做獨立樣本 t 檢定。

$H_0$ ：有得過書卷獎和沒有得過書卷獎的受訪者對於就讀研究所的意願沒有顯著差異

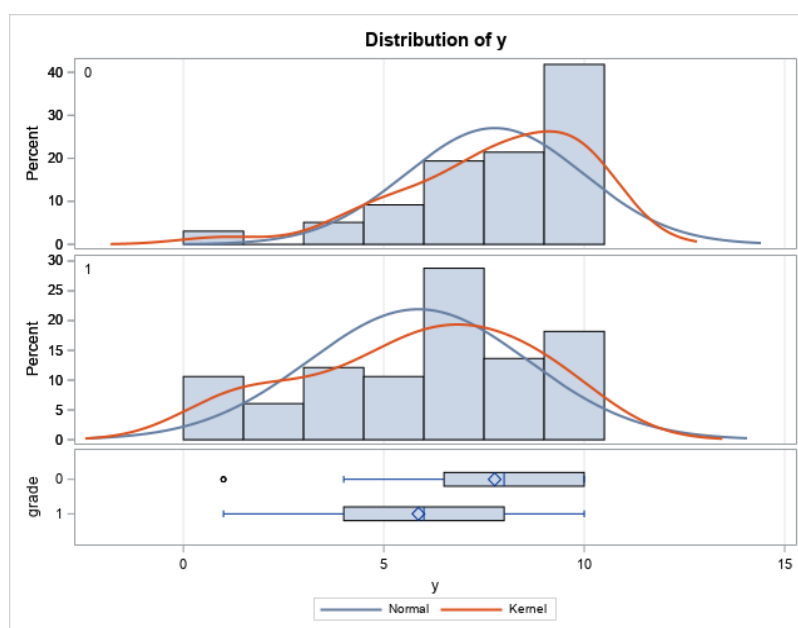
$H_a$ ：有得過書卷獎和沒有得過書卷獎的受訪者對於就讀研究所的意願有顯著差異

經過分析表(七)的結果將如圖(十二)所示。在 Equality of Variances 表中， $p - value$  為 0.1893，明顯大於顯著水準  $\alpha = 0.05$ ，所以要參考 Pooled 方法的  $p - value$ 。在表(七)中，Pooled 的  $p - value$  為 0.0101，明顯小於顯著水準  $\alpha = 0.05$ 。則拒絕  $H_0$  虛無假設，說明有得過書卷獎和沒有得過書卷獎的受訪者對於就讀研究所的意願有顯著差異。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

表(八)成績排名與就讀研究所意願的獨立 t 檢定

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	162	4.90	<.0001
Satterthwaite	Unequal	119.53	4.71	<.0001
Method	Num DF	Den DF	F Value	Pr > F
Folded F	134	65	97	1.52



圖(十三)成績排名與就讀研究所意願的獨立 t 檢定

兩組樣本分別是工科院受訪者中，班排名在 1%到 40%和班排名在 41%到 100%的學生對於就讀研究所的意願程度。我們想要探究這兩組樣本有無顯著差異，所以對這兩個樣本做獨立樣本 t 檢定。

$H_0$ ：班排名在 1%到 40%和班排名在 41%到 100%的受訪者對於就讀研究所的意願沒有顯著差異

$H_a$ ：班排名在 1%到 40%和班排名在 41%到 100%的受訪者對於就讀研究所的意願有顯著差異

經過分析表(八)的結果將如圖(十三)所示。在 Equality of Variances 表中， $p - value$ 為 0.0595，大於顯著水準 $\alpha = 0.05$ ，所以要參考 Pooled 方法的 $p - value$ 。在表(八)中，Pooled 的 $p - value$ 小於 0.0001，明顯小於顯著水準 $\alpha = 0.05$ 。則拒絕 $H_0$ 虛無假設，說明班排名在 1%到 40%和班排名在 41%到 100%的受訪者對於就讀研究所的意願有顯著差異。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

透過四個獨立樣本 t 檢定，我們可以發現：逢甲大學工科院學生對於研究所畢業後的起薪，無論是男生或女生、大三或大四的學生，他們認為研究所畢業後的起薪都沒有顯著差異。說明性別和年級並不會影響逢甲大學工科院學生對於研究所畢業後的起薪認知。逢甲大學工科院學生對於就讀研究所意願，有獲得過書卷獎的學生比沒有獲得過書卷獎的學生、班排名在前 1%到 40%的學生比班排名在 41%到的學生，更願意就讀研究所。



## 第四章、迴歸分析

### 第一節、變數選取

此研究母體總人數為 2164 人選取 165 人為樣本。共有 46 個變數包含 19 個連續型變數以及 27 個虛擬變數。進行迴歸分析前，首先需要進行變數選取，透過變異數膨脹因子(VIF)檢查自變數之間是否有多重共線性。再以逐步選取法、向前選取法以及向後消去法三種方式，分別探討選取的重要變數。

#### (一)、多重共線性

表(九)各變數的變異數膨脹因子數值

Variable	Intercept	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
VIF	0.00	2.57	2.25	2.93	3.32	2.43	2.24	1.80	1.87	3.54	2.36	3.14
Variable	x12	x13	x14	x15	x16	x17	x18	D19	x20	x21	x22	x23
VIF	3.25	3.28	4.39	3.67	2.69	2.27	3.08	1.81	2.54	2.16	1.72	2.46
Variable	x24	x25	x26	x27	x28	x29	x30	x31	x32	x33	D34	x35
VIF	2.22	1.53	2.01	2.06	2.66	3.44	2.17	1.71	1.82	4.50	1.47	1.93
Variable	D1	D2	D3	D4	D5	D6	D7	D8	D9	x37	x38	
VIF	5.16	3.99	4.00	4.11	3.74	4.39	4.06	4.04	3.71	1.35	1.52	

迴歸模型中，自變數之間存在高度相關，可能會造成多重共線性的原因可能有以下幾點：

- 1、解釋變數之間在時間上有共同的趨勢
- 2、由於資料樣本數不夠多，有些解釋變數之間可能會互相影響
- 3、解釋變數之間可能存在相似的線性關係

可以利用變異數膨脹因子(VIF)判斷模型有無多重共線性。一般來說，VIF 介於 0 到 10 之間代表沒有多重共線性，VIF 大於 10 則代表有多重共線性。

從表(九)中 VIF 數值可以發現，沒有任何解釋變數造成多重共線性。因此將選取所有變數進行迴歸選取。

## (二)、選取法

表(十)迴歸變數選取法

Model Selection	Variables	Number of steps	Typical Criteria
Stepwise Selection	x1 x4 x5 x6 x9 x17 x18 x21 x32 x33 D34	11	Entry : $p - value < 0.1$ Remove : $p - value > 0.15$
Forward Selection	x1 x4 x5 x6 x9 x17 x18 x21 x32 x33 D34	11	Entry : $p - value < 0.1$
Backward Elimination	x1 x4 x5 x11 x12 x17 x18 x22 x23 x32 x33 D34	26	Remove : $p - value > 0.1$

逐步選取法的優勢：逐步選取法具備更嚴謹的自變數篩選機制，能夠避免不合適的自變數選入迴歸模型中，進而影響最終的解釋程度及預測正確性。

篩選機制：自變數在每一步驟的篩選時，都有訂定挑選與剔除的標準值。因此逐步選取法較向前、向後消去法更加嚴謹，能夠確保在整個過程中只挑選有顯著影響的變數，並且重複篩選每個步驟所挑選的變數均對於新配適的迴歸模型都具有顯著影響。

因此重要變數的選取法採用逐步選取法，設定挑選變數的顯著值為小於 0.10，而剔除變數的顯著值為大於 0.15，共進行 11 次的篩選，其中有 x1、x4、x5、x6、x9、x17、x18、x21、x32、x33 與 D34 共 11 個變數為迴歸模型的自變數。

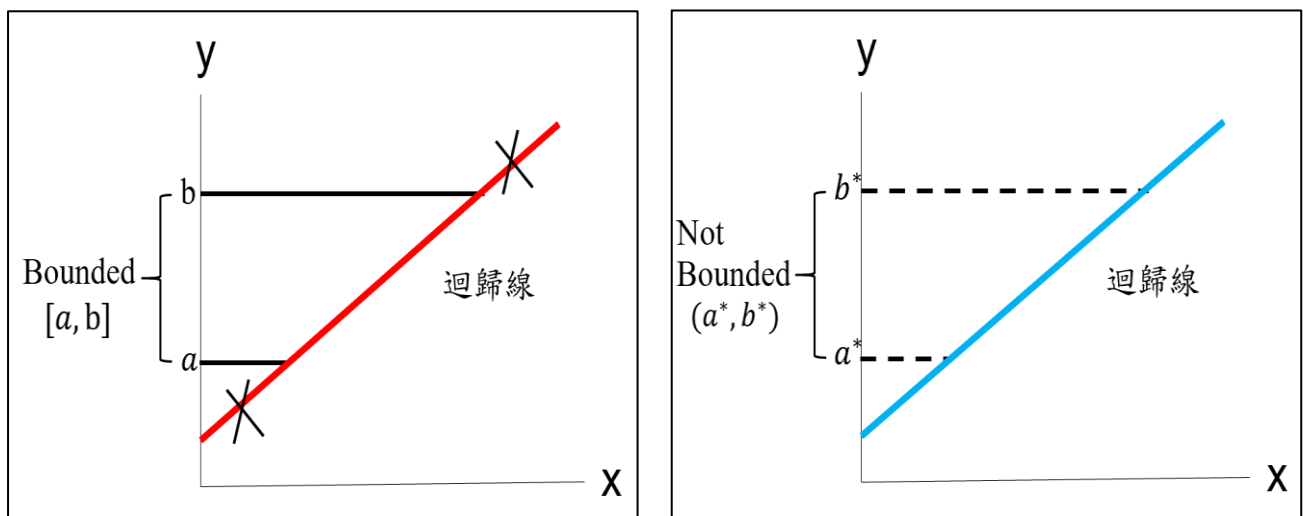
## (三)、有界變數轉換(Bounded Variable Transformation)

我們利用 Bounded Variable Transformation 進行變數轉換，其轉換公式如下：

$$yq = \ln \left( \frac{y - \min + \varepsilon}{\max - y + \varepsilon} \right) * c, \text{ 其中 } \varepsilon \text{ 為誤差, } c \text{ 為常數。}$$

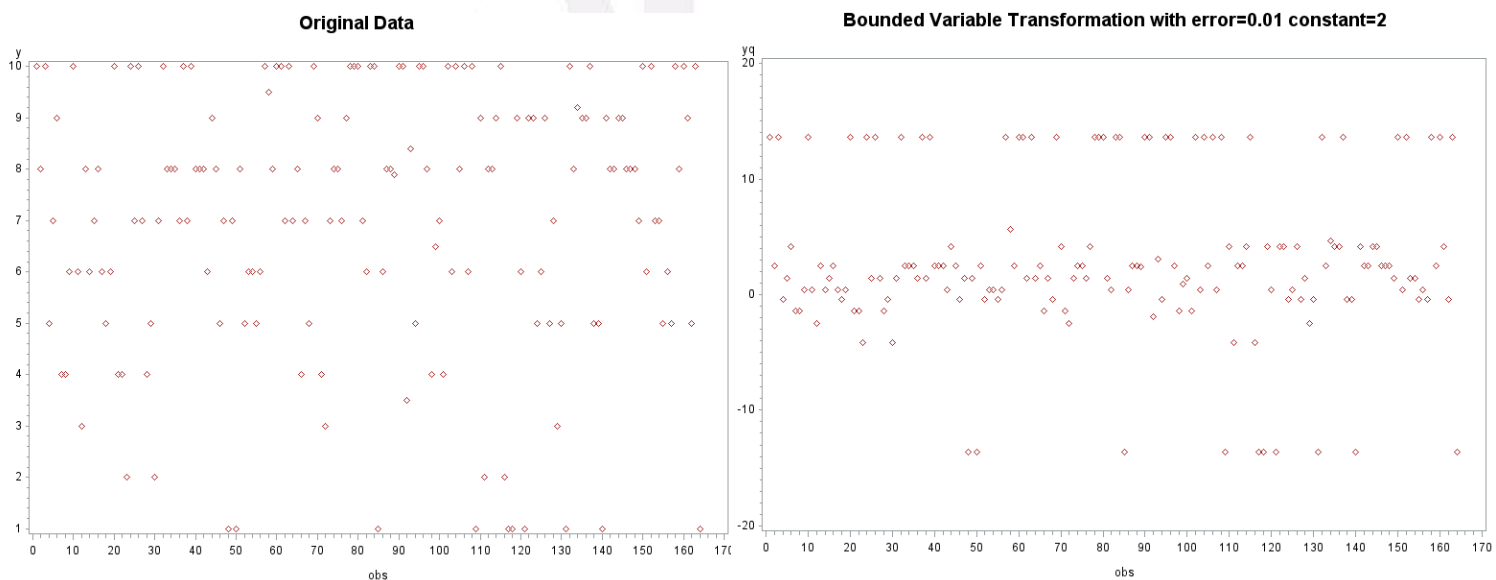
其轉換主要目的為，使連續型變數  $y$  的範圍從閉區間  $y = [a, b]$  圖(十四 a)，轉變為開區間  $yq = (a^*, b^*)$  圖(十四 b)，進而使得超出變數  $y$  所定義之範圍的預測值  $\hat{y}$  有意義。由於變數  $y$  代表就讀研究所的意願程度，1 代表意願程度最低，10 代表意願程度最高，因此必須經由變數轉換，使得區間從原本被束縛的閉區間  $y = [1, 10]$  圖(十五 a)，轉變為開區間  $yq = (-13.607, 13.607)$  圖(十五 b)。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)



圖(十四)假設連續型變數  $y$  的範圍

(a)：閉區間 (b)開區間



圖(十五)實際連續型變數  $y$  的範圍

(a)：閉區間 (b)開區間

經過變數轉換後，可以由圖(十五)觀察出區間的變化，從原本被束縛的閉區間： $y = [1, 10]$ ，轉變為開區間： $yq = (-13.607, 13.607)$ ， $\varepsilon = 0.01$ ， $c = 2$ 。



## 第二節、配適模型

我們透過 SAS 針對 46 個變數進行迴歸分析，配適 5 個模型。根據表(十一)的結果，比較變異數分析所得到的 F 值、R-square、調整 R-square、 $\beta$  係數之 t 值、殘差自我相關檢定與殘差常態分布檢定等變化，可以發現 Final Model 的結果最為理想。

表(十一)迴歸分析結果

迴歸模型	觀察值數量	應變數	F (p-value)	R <sup>2</sup>	Adjusted R <sup>2</sup>	迴歸係數 t-test	Durbin-Watson 獨立性檢定	Shapiro-Wilk 常態檢定
Full Model	161	y	9.33 (<.0001)	0.7901	0.7054	X	2.2832 ✓	0.9906 ✓
1	164	yq	29.59 (<.0001)	0.6817	0.6587	X	2.010 ✓	0.9826 X
2	164	yq	31.85 (<.0001)	0.6755	0.6543	X	2.0527 ✓	0.9775 X
3	163	yq	35.43 (<.0001)	0.6998	0.6800	✓	1.950 ✓	0.9921 ✓
Final Model	162	yq	38.51 (<.0001)	0.7183	0.6997	✓	2.0443 ✓	0.9956 ✓

✓：表示檢定通過 X：表示檢定不通過

*Full Model*：共 46 個變數包含 19 個連續型變數及 27 個虛擬變數

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_7 x_{7i} + \beta_8 x_{8i} + \beta_9 x_{9i} + \beta_{10} x_{10i} + \beta_{11} x_{11i} + \beta_{12} x_{12i} \\ + \beta_{13} x_{13i} + \beta_{14} x_{14i} + \beta_{15} x_{15i} + \beta_{16} x_{16i} + \beta_{17} x_{17i} + \beta_{18} x_{18i} + \beta_{19} D_{19i} + \beta_{20} x_{20i} + \beta_{21} x_{21i} + \beta_{22} x_{22i} + \beta_{23} x_{23i} \\ + \beta_{24} x_{24i} + \beta_{25} x_{25i} + \beta_{26} x_{26i} + \beta_{27} x_{27i} + \beta_{28} x_{28i} + \beta_{29} x_{29i} + \beta_{30} x_{30i} + \beta_{31} x_{31i} + \beta_{32} x_{32i} + \beta_{33} x_{33i} + \beta_{34} D_{34i} \\ + \beta_{35} x_{35i} + \beta_{36} D_{1i} + \beta_{37} D_{2i} + \beta_{38} D_{3i} + \beta_{39} D_{4i} + \beta_{40} D_{5i} + \beta_{41} D_{6i} + \beta_{42} D_{7i} + \beta_{43} D_{8i} + \beta_{44} D_{9i} + \beta_{45} x_{37i} + \beta_{46} x_{38i} + \varepsilon_i$$

(1) 觀察值數量：161 筆

(2) 觀察值處理：刪除第 11 筆、第 65 筆、第 151 筆與第 158 筆觀察值。

第 11 筆觀察值：受訪者未按照問卷第一部份填寫格式。受訪者於變數 x12、變數 x13 與變數 x14 填寫 0，與問卷格式不符。

第 66 筆觀察值：變數 x35 有缺失值。

第 152 筆觀察值：變數 x35 有缺失值。

第 159 筆觀察值：變數 x35 有缺失值。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎?以逢甲工科院為例)

- (3) 變數轉換：未進行變數轉換，直接令  $y$  為迴歸模型的應變數。
- (4) 分析方法及檢定：迴歸分析、殘差診斷分析。
- (5) 重要變數： $x_2$ 、 $x_3$ 、 $x_4$ 、 $x_9$ 、 $x_{11}$ 、 $x_{12}$ 、 $x_{17}$ 、 $x_{18}$ 、 $D_{19}$ 、 $x_{21}$ 、 $x_{26}$
- (6) 分析結果：此迴歸模型對預測結果的可解釋程度約 79%，但部分自變數之 $\beta$ 係數不具有解釋力。殘差之間互相獨立且殘差服從常態分布。

Model 1：共 11 個變數包含 7 個連續型變數及 4 個虛擬變數

$$yq_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_9 x_{9i} + \beta_{17} x_{17i} + \beta_{18} x_{18i} + \beta_{21} x_{21i} + \beta_{32} x_{32i} + \beta_{33} x_{33i} + \beta_{34} D_{34i} + \varepsilon_i$$

- (1) 觀察值數量：164 筆
- (1) 觀察值處理：刪除第 11 筆觀察值。  
第 11 筆觀察值：受訪者未按照問卷第一部份填寫格式
- (2) 分析方法及檢定：逐步迴歸法、迴歸分析、殘差診斷分析。
- (3) 重要變數： $x_1$ 、 $x_4$ 、 $x_5$ 、 $x_6$ 、 $x_9$ 、 $x_{17}$ 、 $x_{18}$ 、 $x_{21}$ 、 $x_{32}$ 、 $x_{33}$ 、 $D_{34}$
- (4) 分析結果：此迴歸模型對預測結果的可解釋程度約 66%，但部分自變數之 $\beta$ 係數不具有解釋力。殘差之間互相獨立但殘差不服從常態。

Model 2：共 10 個變數包含 6 個連續型變數及 4 個虛擬變數

$$yq_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_9 x_{9i} + \beta_{18} x_{18i} + \beta_{21} x_{21i} + \beta_{32} x_{32i} + \beta_{33} x_{33i} + \beta_{34} D_{34i} + \varepsilon_i$$

- (2) 觀察值數量：164 筆
- (3) 觀察值處理：刪除第 11 筆觀察值。  
第 11 筆觀察值：受訪者未按照問卷第一部份填寫格式
- (4) 變數處理：因逐步迴歸選取的自變數  $x_{17}$  的 $\beta$ 係數不具有解釋力，所以將其刪除。
- (5) 變數轉換： $y$  進行變數轉換為  $yq$ ，其  $\varepsilon=0.01$ 、 $c=2$ ，令  $yq$  為迴歸模型的應變數。
- (6) 分析方法及檢定：逐步選取法、迴歸分析、殘差自我相關檢定、殘差常態分布檢定。
- (7) 重要變數： $x_1$ 、 $x_4$ 、 $x_5$ 、 $x_6$ 、 $x_9$ 、 $x_{18}$ 、 $x_{21}$ 、 $x_{32}$ 、 $x_{33}$ 、 $D_{34}$
- (8) 分析結果：此迴歸模型對預測結果的可解釋程度約 68%，但部分自變數之 $\beta$ 係數不具有解釋力。殘差之間互相獨立但殘差不服從常態。

*Model 3*：共 10 個變數包含 6 個連續型變數及 4 個虛擬變數

$$yq_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_9 x_{9i} + \beta_{18} x_{18i} + \beta_{21} x_{21i} + \beta_{32} x_{32i} + \beta_{33} x_{33i} + \beta_{34} D_{34i} + \varepsilon_i$$

(1) 觀察值數量: 163 筆

(2) 觀察值處理: 刪除第 11 筆與第 141 筆觀察值。

第 11 筆觀察值：受訪者未按照問卷第一部份填寫格式。

第 141 筆觀察值：其標準化殘差為-4.072，視為嚴重異常點。

(3) 變數處理：因逐步迴歸選取的自變數 x17 的  $\beta$  係數不具有解釋力，所以將其刪除。

(4) 變數轉換：y 進行變數轉換為 yq，其  $\varepsilon=0.01$ 、 $c=2$ ，令 yq 為迴歸模型的應變數。

(5) 分析方法及檢定：逐步迴歸法、迴歸分析、殘差診斷分析、異常點診斷。

(6) 重要變數：x1、x4、x5、x6、x9、x18、x21、x32、x33、D34

(7) 分析結果：此迴歸模型對預測結果的可解釋程度約 70%，自變數之  $\beta$  係數均具有解釋力，殘差之間互相獨立且殘差服從常態分布。

*Final Model*：共 10 個變數包含 6 個連續型變數及 4 個虛擬變數

$$yq_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_9 x_{9i} + \beta_{18} x_{18i} + \beta_{21} x_{21i} + \beta_{32} x_{32i} + \beta_{33} x_{33i} + \beta_{34} D_{34i} + \varepsilon_i$$

(1) 觀察值數量：162 筆

(2) 觀察值處理：刪除第 11 筆、第 141 筆與第 96 筆觀察值。

第 11 筆觀察值：受訪者未按照問卷第一部份填寫格式。

第 141 筆觀察值：其標準化殘差為-4.072，視為嚴重異常點。

第 96 筆觀察值：其標準化殘差為 3.406，視為嚴重異常點。

(3) 變數處理：因逐步迴歸選取的自變數 x17 的  $\beta$  係數不具有解釋力，所以將其刪除。

(4) 變數轉換：y 進行變數轉換為 yq，其  $\varepsilon = 0.01$ 、 $c=2$ ，令 yq 為迴歸模型的應變數。

(5) 分析方法及檢定：逐步迴歸法、迴歸分析、殘差診斷分析、異常點診斷。

(6) 重要變數：x1、x4、x5、x6、x9、x18、x21、x32、x33、D34

(7) 分析結果：此迴歸模型對預測結果的可解釋程度約 72%，自變數之  $\beta$  係數均具有解釋力，殘差之間互相獨立且殘差服從常態分布。

### 第三節、最終模型

此章節的目的要探前一章節的第五個模型 Final Model，在經過逐步迴歸法選取重要變數；刪除自變數 x17；診斷分析刪除 2 個異常點後，了解模型的配適結果以及殘差是否符合殘差的三大基本假設，以支持模型的預測結果不會有所偏誤。

*Final Model*：共 10 個變數包含 6 個連續型變數及 4 個虛擬變數

$$yq_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_9 x_{9i} + \beta_{18} x_{18i} + \beta_{21} x_{21i} + \beta_{32} x_{32i} + \beta_{33} x_{33i} + \beta_{34} D_{34i} + \varepsilon_i$$

表(十二)*Final Model* 的 ANOVA 表

Analysis of Variance					
Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Model	10	5168.075	516.8075	38.51	<.0001
Error	151	2026.385	13.41977		
Corrected Total	161	7194.461			

表(十三)*Final Model* 的參數估計表

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variance Inflation
Intercept	1	-16.15647	1.74047	-9.28	<.0001	0
x1	1	0.59135	0.17265	3.43	0.0008	1.73342
x4	1	0.86400	0.19686	4.39	<.0001	1.90836
x5	1	0.70900	0.21906	3.24	0.0015	1.72014
x6	1	-0.65547	0.20752	-3.16	0.0019	1.69390
x9	1	0.66914	0.21413	3.12	0.0021	2.11937
x18	1	0.77758	0.16993	4.58	<.0001	1.76076
x21	1	-1.84699	0.76605	-2.41	0.0171	1.06892
x32	1	-3.08583	1.39495	-2.21	0.0285	1.10273
x33	1	-2.33128	0.63323	-3.68	0.0003	1.19833
D34	1	1.92264	0.77212	2.49	0.0139	1.05770

由表(十二)可得知，在刪除逐步迴歸法的 11 個重要變數的其中一個自變數 x17，及該模型的兩個 2 個異常點後，F 統計量的 *p-value* 小於 5% 顯著水準，因此拒絕虛無假設，故表示該模型的自變數與應變數具有線性相關。

(人生轉捩點-大學畢業後要繼續就讀研究所嗎?以逢甲工科院為例)

由表(十三)可得知，在刪除逐步迴歸法的 11 個重要變數的其中一個自變數  $x_{17}$ ，及該模型的兩個異常點後，所有自變數  $t$  統計量的  $p$ -value 均小於 5% 顯著水準，因此拒絕虛無假設，故表示該模型的自變數  $\beta$  係數與應變數具有線性相關。

在執行迴歸分析的  $F$  檢定與  $t$  檢定前，需檢測模型的殘差是否符合三大基本假設。若資料的殘差有違反基本假設，將會造成統計推論的偏誤，因此模型的預測結果會不精確。其中殘差三大基本假設如下：

$$\begin{aligned} E(\varepsilon_i) &= 0 \\ \text{Cov}(\varepsilon_i, \varepsilon_j) &= 0, \forall i \neq j \\ \varepsilon_i &\sim N(0, \sigma^2) \end{aligned}$$

### (一)、檢定殘差的平均數為零

$$H_0: E(\varepsilon_i) = 0$$

$$H_a: E(\varepsilon_i) \neq 0$$

表(十四)殘差的平均數檢定

Tests for Location: $\mu_0=0$				
Test	Statistic		p Value	
Student's t	t	0	$\text{Pr} >  t $	1.0000
Sign	M	-2	$\text{Pr} \geq  M $	0.8138
Signed Rank	S	60.5	$\text{Pr} \geq  S $	0.9198

由表(十四)殘差的 Location Tests 得知，Student's t、Sign、Signed Rank 的  $p$ -value 均大於 5% 顯著水準，因此不拒絕虛無假設，故表示殘差的平均數為 0。

### (二)、檢定殘差之間是否獨立 $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

正自我相關假設

$$H_0: \rho = 0$$

$$H_a: \rho > 0$$

負自我相關假設

$$H_0: \rho = 0$$

$$H_a: \rho < 0$$

表(十五)檢驗殘差的自我相關

Durbin-Watson Statistics			
Order	DW	$\text{Pr} < \text{DW}$	$\text{Pr} > \text{DW}$
1	2.0443	0.6037	0.3963

(人生轉捩點-大學畢業後要繼續就讀研究所嗎？以逢甲工科院為例)

由表(十五)殘差的 Durbin-Watson 檢定得知，DW 的  $p$ -value 均大於 5% 顯著水準，因此不拒絕虛無假設，故表示殘差之間沒有正自我相關及負自我相關。

### (三)、檢定殘差是否服從常態分布

$H_0$ : 殘差服從常態分布

$H_a$ : 殘差不服從常態分布

表(十六)殘差的常態性檢定

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.995559	Pr < W	0.9119
Kolmogorov-Smirnov	D	0.031081	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.028304	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.200935	Pr > A-Sq	>0.2500

由表(十六)殘差的常態檢定得知，Shapiro-Wilk、Kolmogorov-Smirnov、Cramer-von Mises、Anderson-Darling 的  $p$ -value 均大於 5% 顯著水準，因此不拒絕虛無假設，故表示殘差服從常態分布。



#### 第四節、迴歸方程式

$$\hat{yq} = -16.156 + 0.5914x_1 + 0.864x_4 + 0.709x_5 - 0.6555x_6 + 0.6691x_9 \\ + 0.7776x_{18} - 1.847x_{21} - 3.0858x_{32} - 2.3313x_{33} + 1.9226D_{34}$$

表(十七)最終迴歸方程式係數

yq	就讀研究所意願程度。	$\beta_0$	-16.156
x1	受 COVID-19 疫情影響，新鮮人求職環境欠佳，我認為就讀研究所是較好的選擇。	$\beta_1$	0.5914
x4	我認為就讀研究所所需的費用支出（例如：學費、生活費）是有價值的。	$\beta_2$	0.864
x5	我認為研究所的學術環境(例如：圖書館、實驗室、研討室)是重要的。	$\beta_5$	0.709
x6	我認為讀研究所能增廣人脈，對未來工作上會有幫助。	$\beta_6$	-0.6555
x9	相較於大學部，我認為研究所能提供更好的實習機會與職業發展前景。	$\beta_9$	0.6691
x18	尚未知道出社會所需之硬實力與軟實力，因此選擇就讀碩士班加強技能。	$\beta_{18}$	0.7776
x21	若我要就讀研究所，我將優先考慮的條件為住家距離。	$\beta_{21}$	-1.847
x32	大學期間，是否參加公務員高、普、特考補習。	$\beta_{32}$	-3.0858
x33	大學期間，是否參加校外學習。	$\beta_{33}$	-2.3313
D34	大學期間，我曾獲得書卷獎。	$\beta_{34}$	1.9226

從最終迴歸方程式可以看出，應變數 yq(就讀研究所意願程度)與自變數 x1、x4、x5、x9、x18、D34 成正相關關係；與自變數 x6、x21、x32、x33 成負相關關係。  
解釋：

- (1) yq 與 x1：受疫情影響，逢甲大學工科院學生選擇就讀研究所意願增強。
- (2) yq 與 x4：認同研究所費用支出是有價值的學生，就讀研究所的意願越強烈。
- (3) yq 與 x5：認同研究所學術環境是重要的學生，就讀研究所的意願越強烈。
- (4) yq 與 x6：認同就讀研究所可以增廣人脈的學生就讀研究所的意願反而下降，原因可能為，直接工作帶來的收益大於就讀研究所的收益。因此即使認同研究所能夠帶來人脈，但依然認為進入職場的效益更大。所以就讀研究所可以增廣人脈的條件並不是增加逢甲大學工科院學生就讀研究所的潛在因素。
- (5) yq 與 x9：認同研究所能夠提供更好實習機會和發展前景的學生，就讀研究所的意願越強烈。
- (6) yq 與 x18：尚未知道社會要求具體硬實力和軟實力的學生，就讀研究所的意願越強烈。
- (7) yq 與 x21：選擇把住家距離作為就讀研究所優先考慮因素的學生，就讀研究所的意願並不強烈。
- (8) yq 與 x32：大學期間選擇參加公務員高、普、特考補習的學生，就讀研究所的意願並不強烈。其可能原因為公務員與研究所考試是生涯發展兩個截然不同的方向。
- (9) yq 與 x33：大學期間選擇不參加任何校外學習的學生，就讀研究所的意願並不強烈。
- (10) yq 與 D34：大學期間有獲得書卷獎的學生，就讀研究所的意願越強烈。



## 第五章、結論

經過多種的敘述統計的圖表呈現、檢定與迴歸分析將資料視覺化，探討逢甲大學工科院大三與大四學生對於就讀研究所的意願程度。

從年級和性別探討就讀研究所的意願程度中，即使工科院的女學生偏少的情況，男生與女生就讀研究所的意願沒有顯著差別，女生的意願程度甚至略優於男生。大四學生對於就讀研究所的意願程度很明確，而大三學生多數還不確定就讀研究所的意願。從系所而論，多數工科院大三與大四學生就讀研究所的意願程度較高，但是環科系大四學生對於就讀研究所的意願程度相對比較低。

就讀研究所優先考慮的因素，從總體而論，工科院的學生對於就讀研究所優先考慮的前三個條件是國(私)立、學校資源以及業界知名度。從研究發現，化工系與材料系的學生較優先考慮學校資源與國(私)立，航太系的學生較優先考慮學校資源、國(私)立與產學合作。工科院由於系所多數是實作科系，比較注重學校資源（如實驗室或軟體等）以及業界的產學合作。

其中對四個變數進行獨立樣本 t 檢定，分別得出的結論是：工科院大三與大四的學生對於研究所畢業後的期望起薪沒有顯著差異；工科院男生和女生對於研究所畢業期望薪資沒有顯著差異；有得過或沒有得過書卷獎的受訪者的就讀研究所的意願有顯著差異；班排名在 1% 到 40% 和班排名在 40% 到 100% 的受訪者的就讀研究所的意願有顯著差異。

歷經多方探討後，多元迴歸分析以及診斷分析的結果當中得出，逢甲大學工科院大三與大四學生對於就讀研究所意願非常積極。就讀研究所意願程度與其他因素而論，雖然受到疫情影響，但逢甲大學工科院學生選擇就讀研究所意願仍非常高。而且越認同 x4 研究所費用支出是有價值的、x5 學術環境是重要的、x9 研究所能夠提供更好實習機會和發展前景、x18 尚未知道社會要求具體硬實力和軟實力以及 D34 大學期間有獲得書卷獎的學生，就讀研究所的意願將越強烈與積極。對於成績越好的學生，就讀研究所的意願程度就越高。

最後得出結論，研究所能夠提供更好實習機會與發展前景或者學校資源充足都會使工科院大三與大四的學生就讀研究所的意願程度提高。這也更能證明，我們所認為的工科院多數以實作為主的想法，實作重視設備與場地，資源的不足會令學生就讀研究所的意願降低。且學生認為就讀研究所付出的花費是具有價值性的。對工科院學生來說，付出的過程得到的不只是學術性內容或實習經驗，還有社會要求具體硬實力和軟實力。多數的工科院學生對於就讀研究所的意願是積極且明確的，我們認為就讀研究所將是一個得到更多經驗與學習的機會。

## 附錄一、參考資料

Abraham, B. and Ledolter, J. (2006) Introduction to regression modeling. Belmont, CA: Duxbury Press.

盒鬚圖. (不詳). 讀取於 2021 年 1 月 11 日, 從

[http://www.stat.nuk.edu.tw/cbme/math/statistic/sta2/s4\\_6/node5.html](http://www.stat.nuk.edu.tw/cbme/math/statistic/sta2/s4_6/node5.html)

散布圖. (2020). 收入 維基百科, 自由的百科全書.

<https://zh.wikipedia.org/w/index.php?title=%E6%95%A3%E5%B8%83%E5%9B%BE&oldid=62566434>

Brown, B. M. (1982). Cramer-von Mises Distributions and Permutation Tests. *Biometrika*, 69(3), 619–624. <https://doi.org/10.2307/2335997>

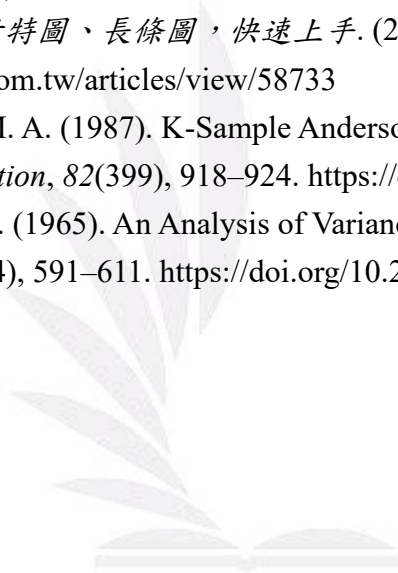
Darling, D. A. (1957). The Kolmogorov-Smirnov, Cramer-von Mises Tests. *The Annals of Mathematical Statistics*, 28(4), 823–838.

Excel 怎麼做圖表？常用甘特圖、長條圖，快速上手. (2019, 十一月 26). 經理人.

<https://www.managertoday.com.tw/articles/view/58733>

Scholz, F. W., & Stephens, M. A. (1987). K-Sample Anderson-Darling Tests. *Journal of the American Statistical Association*, 82(399), 918–924. <https://doi.org/10.2307/2288805>

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), 591–611. <https://doi.org/10.2307/2333709>



(人生轉捩點-大學畢業後要繼續就讀研究所嗎?以逢甲工科院為例)

#### 原創性報告



#### 主要來源

1	Submitted to Feng Chia University 學生文稿	2%
2	sec.nsysu.edu.tw 網際網絡來源	1%
3	www.eurojournals.com 網際網絡來源	1%
4	dspace.lib.fcu.edu.tw 網際網絡來源	<1%
5	www.iecs.fcu.edu.tw 網際網絡來源	<1%
6	www3.grips.ac.jp 網際網絡來源	<1%
7	www.coursehero.com 網際網絡來源	<1%
8	erlang.se 網際網絡來源	<1%
9	www.ba.fcu.edu.tw 網際網絡來源	<1%