

統計學系碩士在職專班

多變量分析

期末報告

111年網路使用與行為調查問卷

指導教授：鍾麗英老師

組別：G06

學生： 711378910林美伶

711378912蔡宜諠

中華民國113年12月

**目錄**

[**第一章 緒論** 3](#_Toc185588025)

[**第一節** **研究背景與動機** 4](#_Toc185588026)

[**第二節** **研究目的** 4](#_Toc185588027)

[**第三節** **研究步驟** 4](#_Toc185588028)

[**第二章 研究方法** 5](#_Toc185588029)

[**第一節** **研究對象** 5](#_Toc185588030)

[**第二節** **問卷設計** 5](#_Toc185588031)

[**第三節** **資料處理** 5](#_Toc185588032)

[**第三章 研究結果** 6](#_Toc185588033)

[**第一節** **敘述性統計** 6](#_Toc185588034)

[**第二節** **主成分分析** 6](#_Toc185588035)

[**第三節** **因素分析** 6](#_Toc185588036)

[**第四節** **集群分析** 6](#_Toc185588037)

[**第五節** **邏輯斯迴歸** 6](#_Toc185588038)

[**第六節** **典型相關** 6](#_Toc185588039)

[**第四章 結論與建議** 7](#_Toc185588040)

[**第一節　結論** 7](#_Toc185588041)

[**第二節　建議** 7](#_Toc185588042)

[**第五章 附錄** 8](#_Toc185588043)

[**第一節　問卷** 8](#_Toc185588044)

[**第二節　問卷編碼** 8](#_Toc185588045)

**第一章 緒論**

1. **研究背景與動機**

經過科技與技術發展，網路已成為許多服務和建設的基礎，也是民眾每日生活的日常，根據財團法人台灣網路資訊中心（TWNIC）《2022台灣網路報告》，目前臺灣網路使用率已超過九成，顯示多數民眾都是網路使用者，無論是日常的食衣住行育樂，醫療就醫、金融等專業領域也都透過網路提供更多便民與個人化服務，疫情期間也透過網路遠端上班、上課，高度利用網路的生活中，民眾透過網路可以取得更多便利，並快速取得許多資訊，然而也可能因此使得人對於網路分離產生焦慮而不可自拔，自己作為網路世代的一員，瞭解網路對身心所可能產生的影響，也能作為警訊，多留意自己的狀況。

透過SRDA學術調查研究資料庫，使用數位發展部委由聯合行銷研究股份有限公司於民國111年12月1日至12月8日執行的「111年網路沉迷研究調查」。研究針對全台12歲以上民眾進行電話訪問調查，主要以「網路使用習慣量表（CIAS-10）」[[1]](#footnote-1)篩選受訪者是否有網路沉迷傾向，問卷中包含網路使用行為資訊，如上網時間、時長、上網活動類型等資料，以及受訪者主觀自評狀況，如網路沉迷自評與心理健康狀況自評，與疫情對生活的影響。

1. **研究目的**

將以研究案主題核心「網路沉迷」出發，運用「網路使用習慣量表（CIAS-10）」進行相關分析的實務操作，以期更清晰探討臺灣民眾網路使用行為，上網者要留意的使用行為關鍵，避免從網路便利轉為依賴再進而沉迷於其中。

需特別說明的是，研究參考衛福部之研究建議，將「有網路沉迷傾向」定義為：在排除工作或課業以外的前提下，CIAS-10 總分達28 分（含）以上的民眾，報告也以此作為是否網路沉迷的變數區分基準。

1. **研究步驟**

透過主成分分析，瞭解在網路使用習慣量表十個題目中，能有助於判別網路沉迷與否的關鍵題目（變數）為何，而藉由因素分析探測網路沉迷分數、網路行為與身心狀況自評等變數的關係，最後透過集群分析將上網使用者進行分群，透過MANOVA分析確認不同分群的上網使用者在哪些上網行為產生顯著差異。

一張含有 文字, 螢幕擷取畫面, Rectangle, 設計 的圖片

自動產生的描述

**第二章 研究方法**

1. **研究對象與調查方法**

本研究以全臺22個縣市（包含臺灣省各縣市、六直轄市及金門、連江縣）年滿12歲且有上網經驗的本國籍民眾為調查對象。採用電腦輔助電話訪問系統（Computer Assisted Telephone Interviewing, CATI）進行電話訪問，使用住宅電話與手機用戶雙底冊調查，蒐集相關調查資料。

1. **抽樣設計**

此問卷參考學術調查研究資料庫，當時採取分層隨機抽樣方式[[2]](#footnote-2)，以全臺22個縣市為分層單位，每縣市視為一個副母體，依據各縣市12歲以上人口比例配置樣本數。樣本抽取過程先以手機號碼隨機抽樣唯手機族，再以住宅電話清冊進行分層隨機抽樣，再以尾數二位隨機變更，使調查結果涵蓋住宅電話未登錄的對象，並反映各縣市實際上網行為的特性。

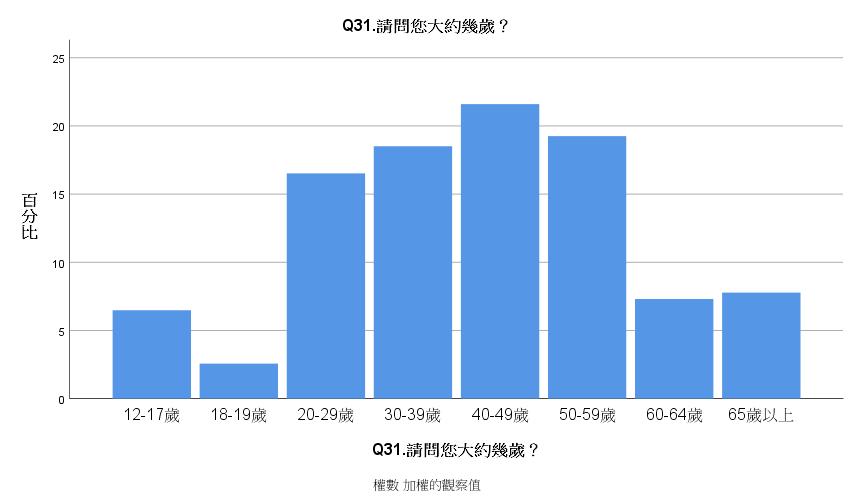
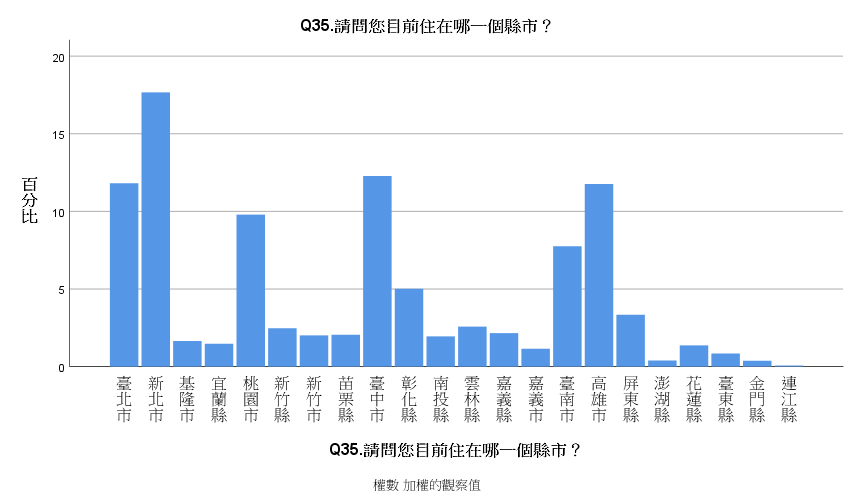
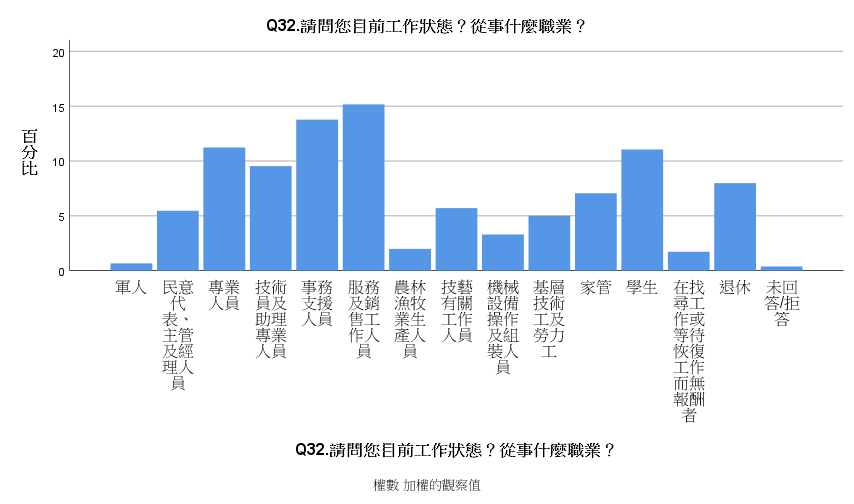
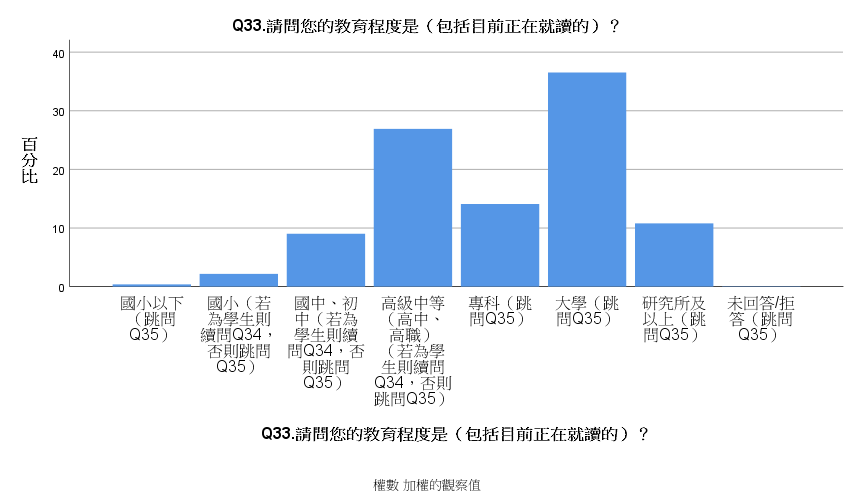
實際完成的有效樣本數為1,975人（含上網及非上網民眾），訪問成功率為68.1%[[3]](#footnote-3)，其中曾上網人數為1,526人。

1. **資料處理**

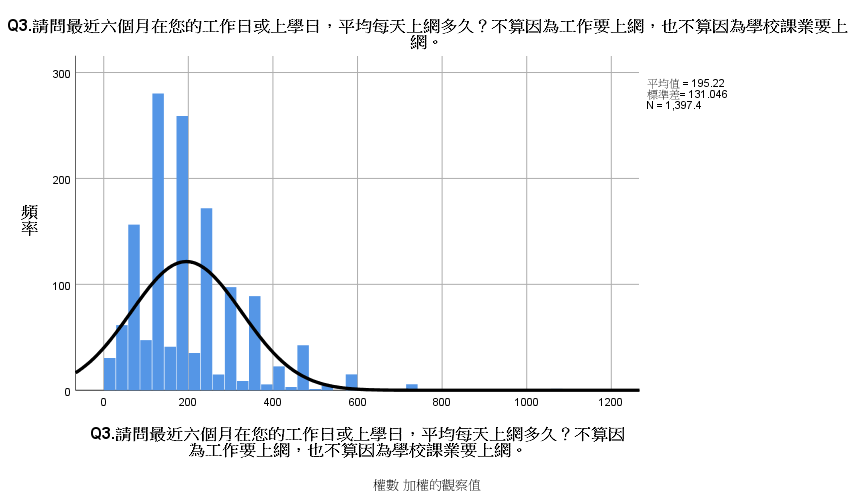
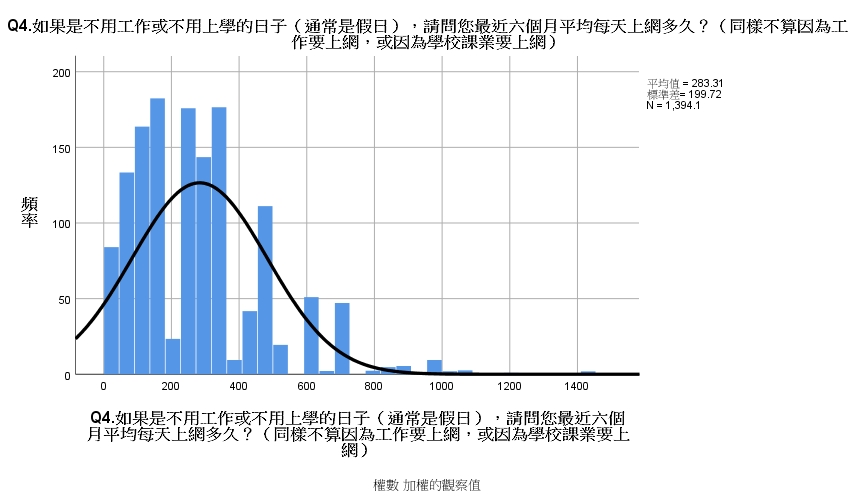
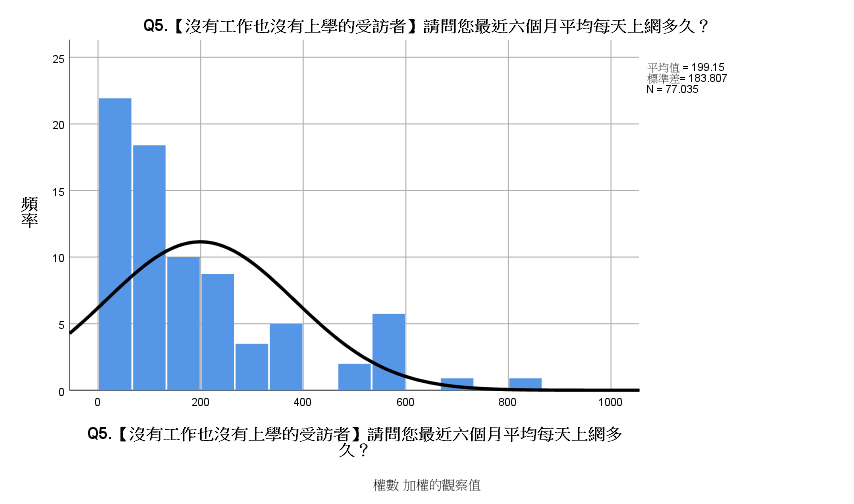
透過SRDA學術調查研究資料庫取得研究資料，以數位發展部提供之資料為基礎，經中央研究院人文社會科學研究中心調查研究專題中心資料整理。原始資料檔案中，已依照臺灣人口比例，依序以性別、年齡、居住縣市進行「多變項反覆加權」，直到每一變項的樣本分配與母體分配的適合度檢定（goodness of fit test）已無顯著差異，並設定每一筆資料的權值（變數wei）。期末報告進行分析時，亦同樣先乘以對應權值後進行分析。

**第三章 研究結果**

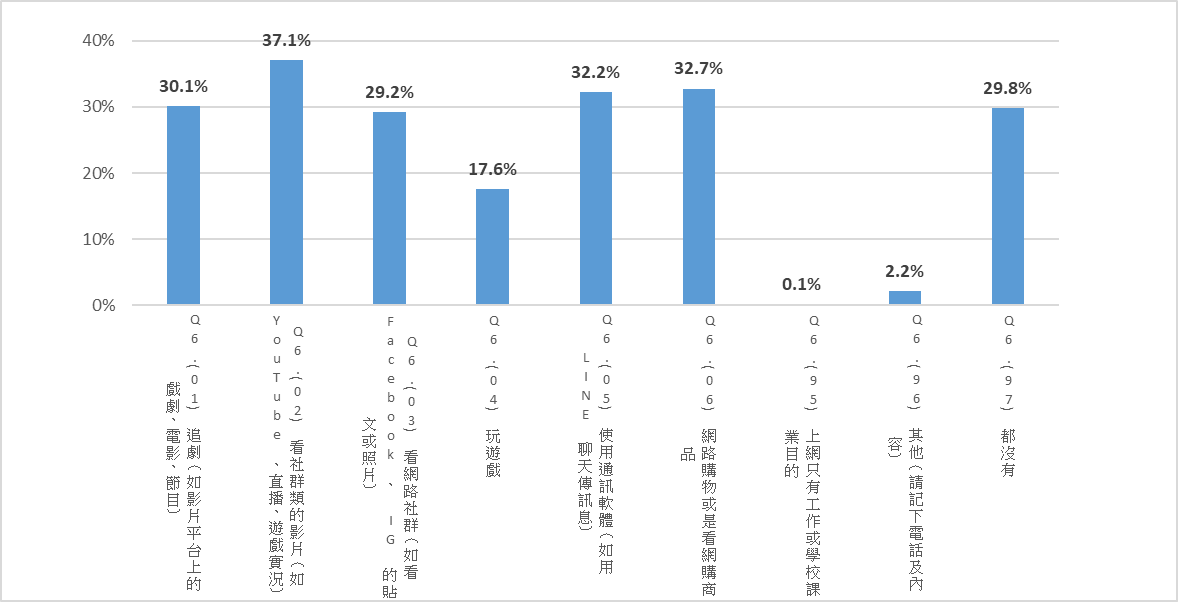
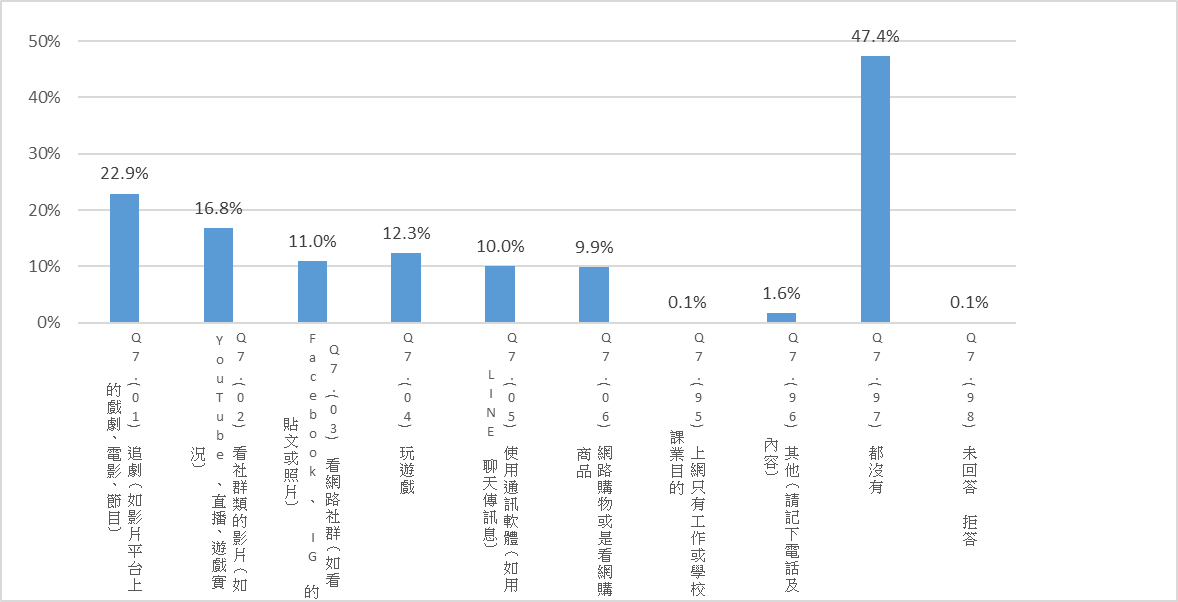
1. **敘述性統計**
2. 基本資料

* 性別：男性和女性受訪者約各半
* 年齡：資料加權後12-19歲受訪者占整體資料9.1%，同臺灣人口分布，以30-59歲受訪者資料占比最高，60-64歲和65歲以上則各占7.3%、7.8%。  
  
* 居住地區（縣市別）：因依照臺灣人口比例性別、地區、年齡反覆加權，加權後受訪者居住地區以六都占比相對較高，其中又以新北市占17.7%為最高，其他臺灣本島縣市則各占1~3%之間，離島的澎湖縣、金門縣和連江縣樣本則皆低於1%，共13位受訪者。  
   
* 職業：受訪者以服務及銷售工作者占比最高（15.2%），其次為事務支援人員（13.8%）。非工作人口（如家管、學生、找尋工作、退休）則占整體樣本的27.8%。  
    
  
* 教育程度：36.5%的受訪者為大學學歷（含就讀中），次多的類別為高中/高職，占整體資料26.9%。  
     
  而針對高中以下的受訪者追問其就讀年級，整體資料中，國小生為10位（0.7%）、國中生44位（2.9%）、高中/高職生49位（3.3%）。  
  

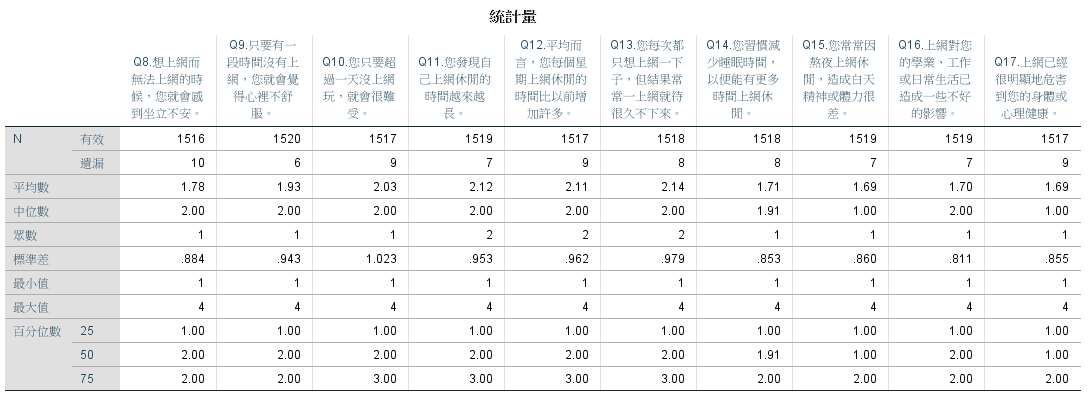
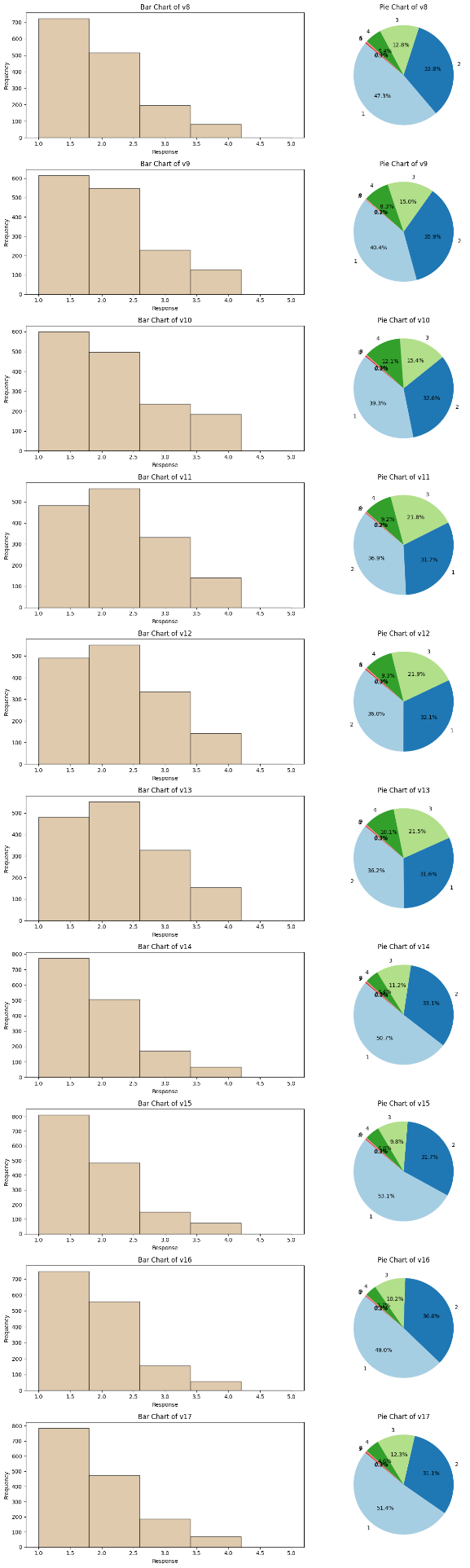
1. 網路使用情形

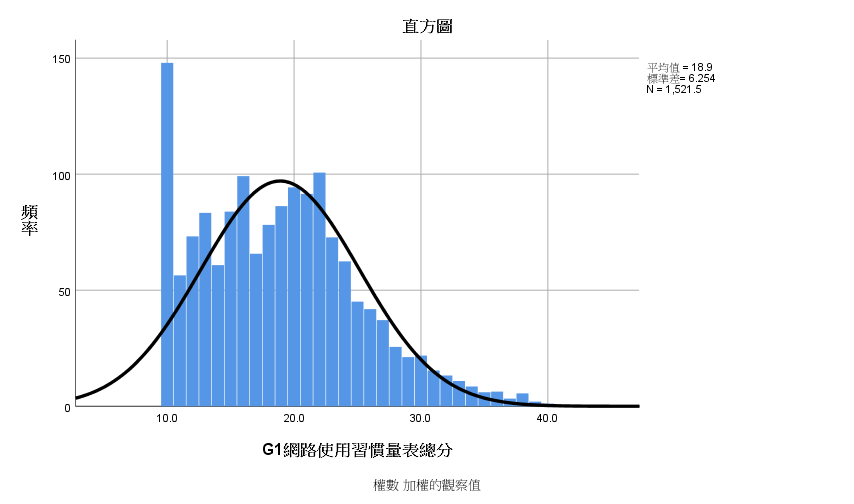
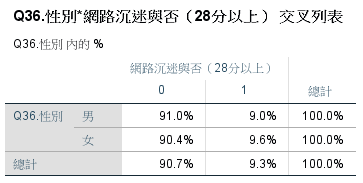
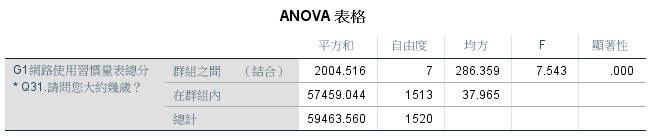
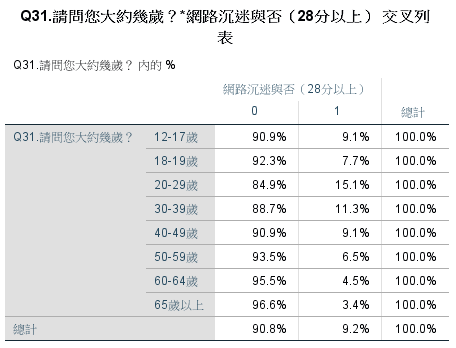
* 上網時數：原始資料為\_\_時\_\_分，經轉換為分鐘數後進行分析。
  + 近六個月工作日/上學日（以下簡稱平日）的非工作/上學上網時數：排除78位近半年沒有工作也沒有上學的受訪者、5位超過一個月沒上網、45位不知道/拒答。針對1397位受訪者的平日非工作/上學上網時數，有0.5%未上網（0分鐘）、6.1%平均上網不到1小時（1~59分鐘）、0.7%平均每天上網超過10小時（600分鐘），其中最多人平日平均上網2小時，占20.1%，其次為上網3小時（18.5%）。整體回答者平日平均上網約3小時15分鐘（195.22分鐘），中位數為3小時。  
    
  + 近六個月非工作日/上學日（以下簡稱休假日）的非工作/上學上網時數：  
    排除近半年沒有工作也沒有上學、近半年沒有休假日和不知道/拒答者，針對1394位受訪者，最多人平均每日上網6小時（11.1%），另有1.1%沒有上網（0分鐘）、5%平均每日上網不到1小時（1~59分鐘）、5.7%平均每天上網超過10小時（600分鐘）。  
    整體回答者平均休假日上網約4小時43分鐘（283.31分鐘），中位數為4小時。  
    
  + 近六個月沒有工作也沒有上學者上網時數：針對78位受訪者進一步詢問其平均上網時數，僅1位回答不知道，針對77位回答者，以平均每日上網2小時的占比最高（19.9%），其次為上網3小時（11.2%）。中位數為平均每日上網2小時，但因為有近一成回答者平均每日上網10小時(含)以上，因此拉高整體平均上網時數至三個多小時（199.15分鐘）。  
    
  + 比較平日與休假日的上網時數，雖兩者皆以日均上網2小時（120分鐘）為多數狀況，但休假日上網時數分布更分散，標準差為199.72，且最大值更來到1440分鐘，中位數提升至240分鐘（4小時），整體而言休假日上網時數更高。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **統計量** | | | | |
|  | | Q3.請問最近六個月在您的工作日或上學日，平均每天上網多久？不算因為工作要上網，也不算因為學校課業要上網。 | Q4.如果是不用工作或不用上學的日子（通常是假日），請問您最近六個月平均每天上網多久？（同樣不算因為工作要上網，或因為學校課業要上網） | Q5.【沒有工作也沒有上學的受訪者】請問您最近六個月平均每天上網多久？ |
| N | 有效 | 1397 | 1394 | 77 |
| 遺漏 | 129 | 132 | 1449 |
| 平均數 | | 195.22 | 283.31 | 199.15 |
| 中位數 | | 180.00 | 240.00 | 120.00 |
| 眾數 | | 120 | 120 | 120 |
| 標準差 | | 131.046 | 199.720 | 183.807 |
| 最小值 | | 0 | 0 | 2 |
| 最大值 | | 1080 | 1440 | 840 |

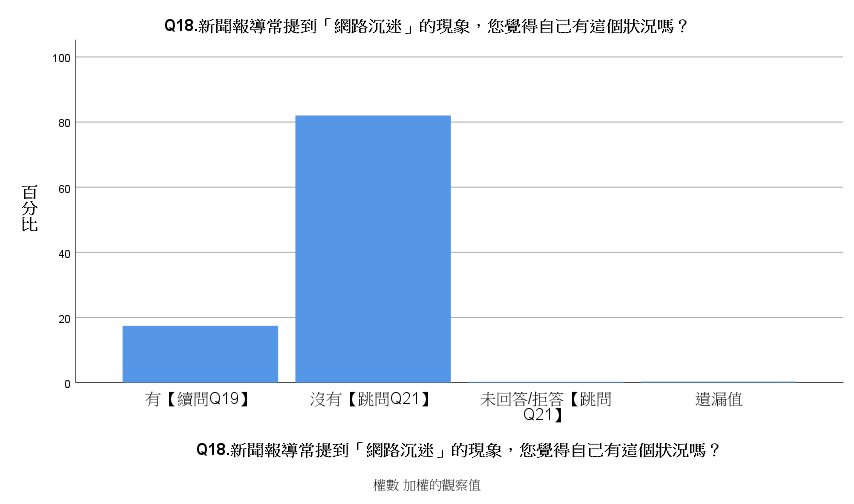
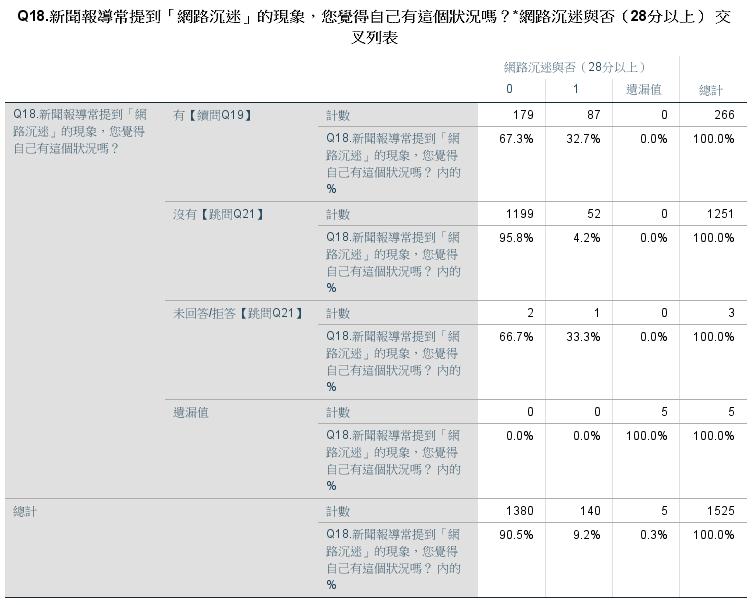
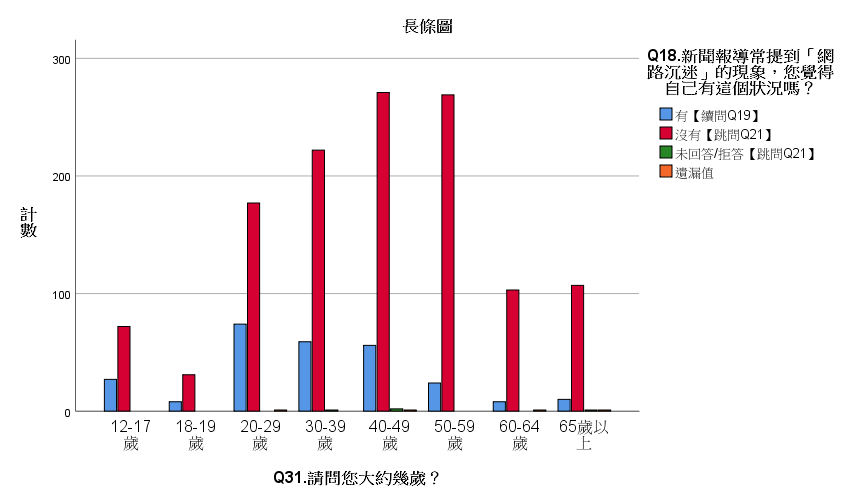
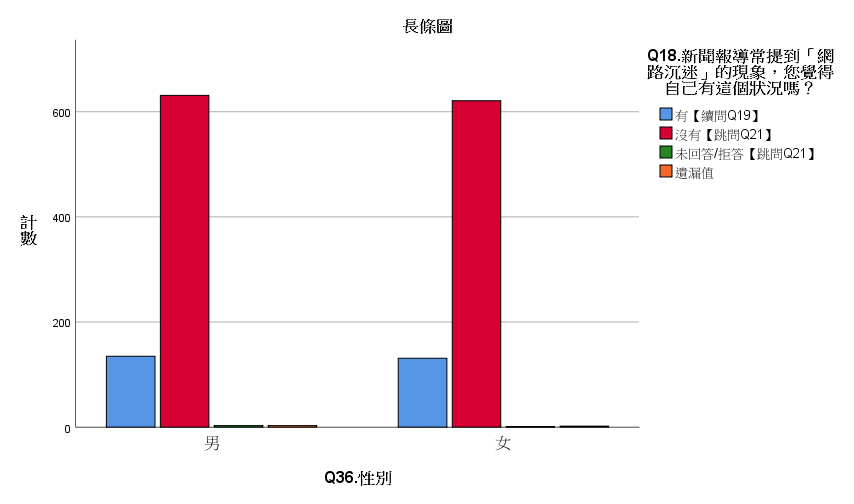
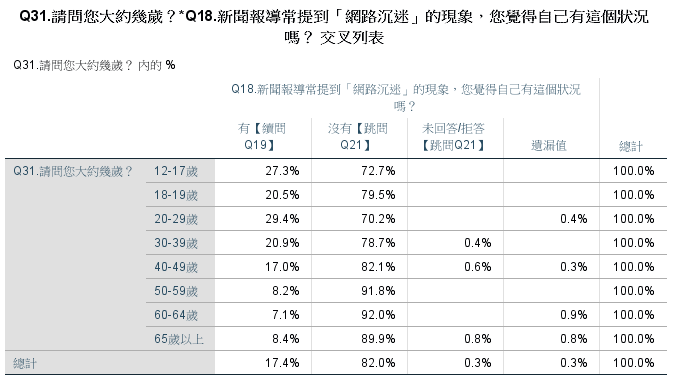
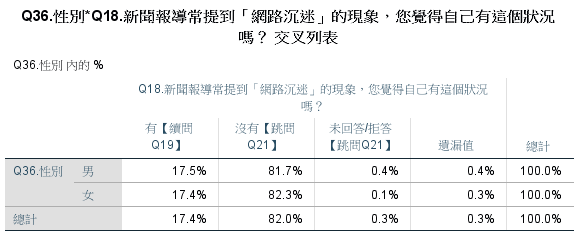
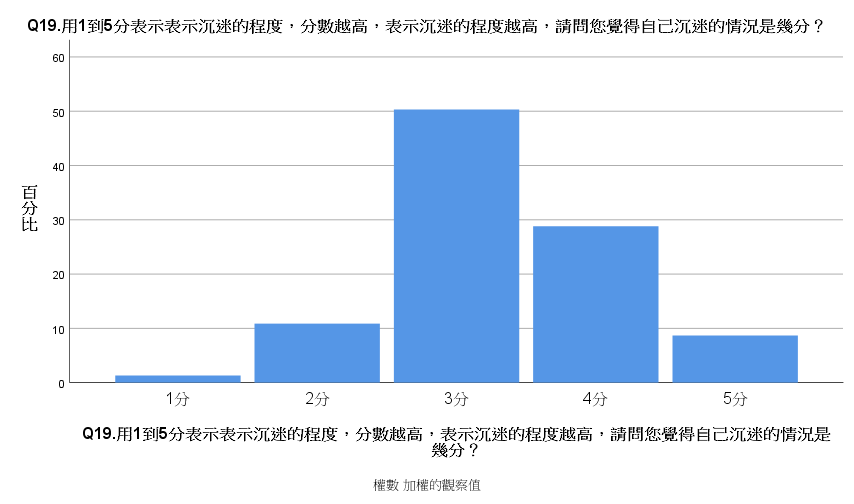
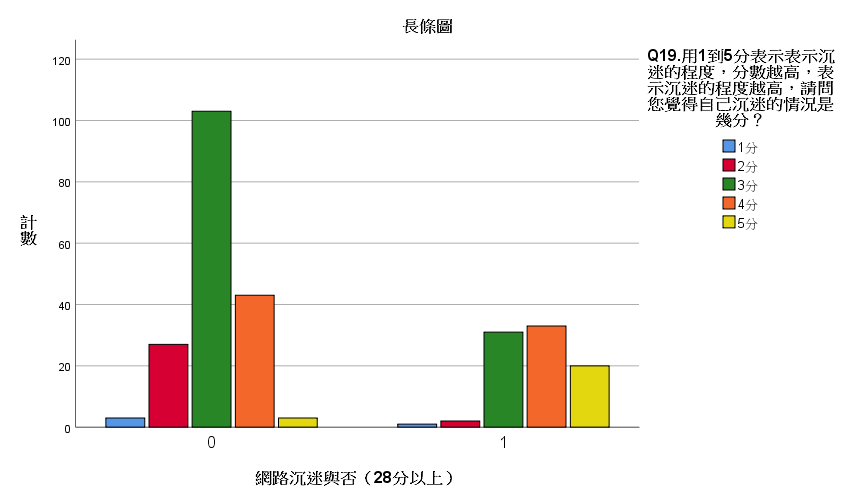
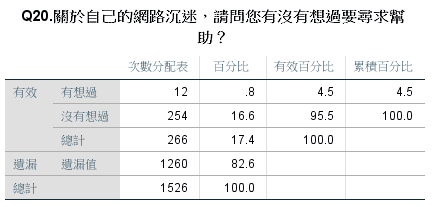
* 與疫情前相比，2022年受訪者自認上網時數不變的比例較多，近七成受訪者認為不變，另有24.2%認為增加，認為減少的比例僅5.9%，但因為本調查上網時數不含工作/上課學習之情境，因此顯示疫情雖有刺激的效果，但對多數民眾日常休閒的上網時數變化不大。  
  
* 與疫情前相比增加投入時間的上網活動：  
  認為沒有因為疫情而增加任何上網活動的受訪者為三成。最多人增加投入時間的上網活動為觀看社群影片，比例37.1%為最高，其次為網路購物和通訊軟體，皆有32%左右的受訪者表示增加網路活動時間。  
  
* 上網無法節制時間的活動類型：  
  從回答顯示約有一半的上網使用者曾於近半年有無法節制時間的上網活動，以追劇的比例最高（22.9%），其次為觀看社群影片（16.8%），電玩遊戲則為12.3%。整體而言，影視類上網活動是最多受訪者疫情後增加投入時間，且相對也有較多比例的使用者難以節制時間。  
  

1. 網路使用習慣量表

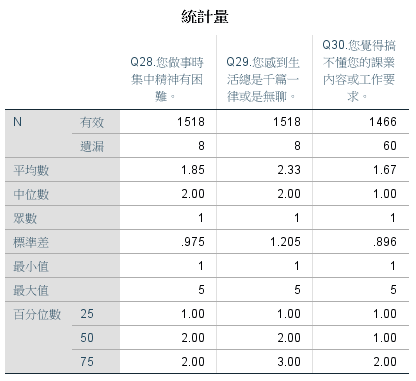
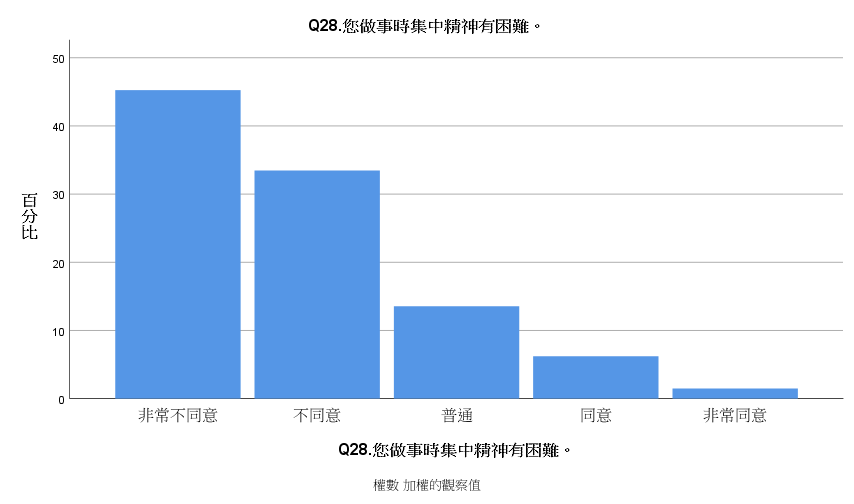
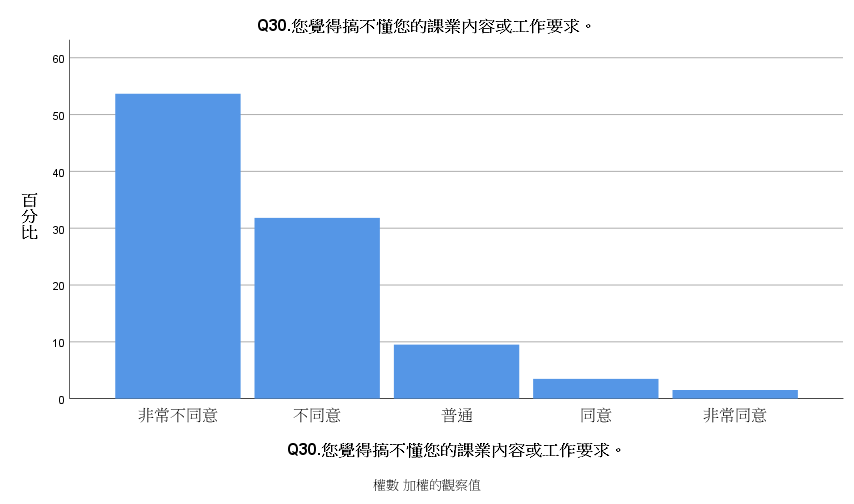
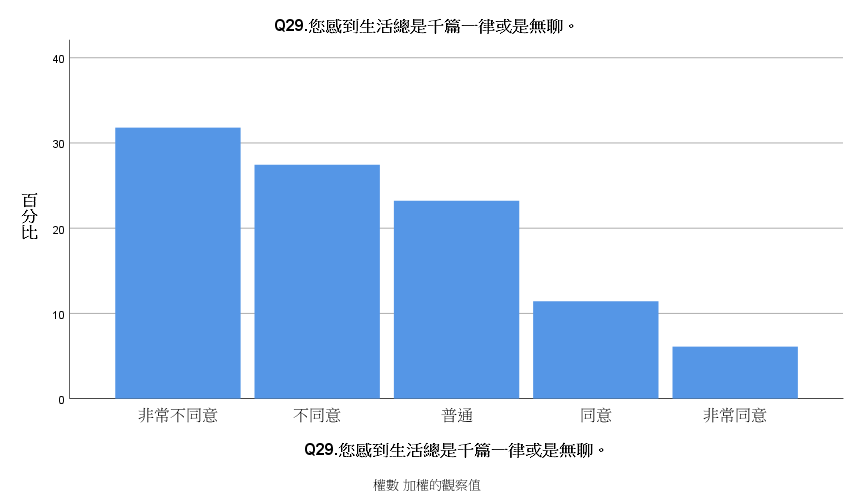
* 量表回答概況：  
  量表以李克特量表四分制進行設置，總題數為10題，1分為非常不符合、2分不符合、3分符合、4分非常符合，題目中並未有反向題。  
  從答題結果來看，「Q15.您常常因熬夜上網休閒，造成白天精神或體力很差」和「Q17.上網已經很明顯地危害到您的身體或心理健康」的平均分數1.69為最低，中位數也為1，超過一半的受訪者沒有這樣的困擾；相反地，「Q10.您只要超過一天沒上網玩，就會很難受」、「Q11.您發現自己上網休閒的時間越來越長」、「Q12.平均而言，您每個星期上網休閒的時間比以前增加許多」、「Q13.您每次都只想上網一下子，但結果常常一上網就待很久不下來」，四個子題的平均分數突破2.0為量表中最高，且中位數提升為2分不符合，第三四分位數更來到3分符合，顯示有25%的受訪者正向表示有此困擾，而從題目敘述來看，主要環繞在上網時間與頻率。  
    
  從子題回答分布可以看到，除了平均分數前三高的Q11、Q12和Q13回答2分布符合的比例最高以外，其他子題皆是分數由低至高逐漸下降的趨勢。  
  一張含有 螢幕擷取畫面, 文字, 設計 的圖片

  自動產生的描述
* 整體量表總分：  
  將網路使用習慣量表（CIAS-10）題項的分數加總，若受訪者所有題項皆為不知道/拒答，則以遺漏值處理，若單一子題未回答，則以該題在量表中具有相同調查目標（因子）的其他子題分數平均值作為插補值，若相同調查目標（因子）的子題皆未作答，則以量表其他題目分數的平均值作為插補值（mean imputation）。  
  整體作答的1521位受訪者，平均分數為18.90分，中位數為19分。從分布圖可看到，以最低分10分的人數最多，占整體回答者9.7%，其次為各占6.5%、6.6%的16分和22分，而第三四分位數為23分，有26.3%回答者的量表總分為23分(含)以上。而符合原研究案的網路沉迷定義則須總分高達28分(含)以上，占回答者9.3%，共140位回答者。  
   
* 各族群是否有網路沉迷：
  + 性別之間的量表總分，首先經F檢定，p值為0.718>.05，因此無法拒絕虛無假設，兩者之間變異數相等，其單一樣本t檢定p值為0..059>0.05，因此性別之間並沒有顯著差異，男性平均總分為18.6、女性平均總分為19.2，男女受訪者間皆各有9%為網路沉迷者。  
      
      
    
  + 年齡層之間量表總分經ANOVA檢定，p值為0.000<.05，表示年齡層之間分數具有顯著差異，其中以20-29歲平均分數最高（20.58），而中位數則隨年齡增長逐漸下降。  
      
      
    而各年齡層屬於網路沉迷者的比例，以20-29歲和30-39歲網路沉迷者占比超過一成最多，其次為40-49歲和12-17歲族群，其中各有9.1%的受訪者為網路沉迷者。  
    

1. 整體網路沉迷主觀自評

* 自評網路沉迷：  
  針對所有受訪者，則有17.4%認為自己有網路沉迷的現象，比例高於原研究案定義下網路沉迷者（網路使用習慣兩表總分28分(含)以上）的比例，9.4%。  
    
  將實際定義與受訪指自評進行交叉分析，實際自評者認為有網路沉迷現象者僅32.7%為實際網路沉迷者，而實際網路沉迷者則有53位沒有病識感，自認沒有網路沉迷的狀況或拒答，占比近四成（53/140=37.86%）。  
    
  而性別之間自認有網路沉迷狀況的比例皆為17%，未有明顯差異，年齡之間則有較明顯的落差，12-39歲皆有超過兩成的受訪者自認有網路沉迷的狀況，又以20-29歲近三成比例最高，50歲以上則不足一成。  
  
* 自評網路沉迷程度：  
  針對266位自認有網路沉迷狀況者，用1~5分表示表示沉迷的程度，分數越高，表示沉迷的程度越高，多數受訪者給予自身3分以上的評價，占比高達86.5%。  
    
  其中實際是網路沉迷者所自評的網路沉迷程度更高，經過單一樣本t檢定，其p值0.000<.05，表示是否為網路沉迷者的網路沉迷自評程度具有顯著差異，網路沉迷者自評程度分數高達3.81，顯著高於非網路沉迷者（平均3.09），自評分數集中於3分以上，相較之下實際不是網路沉迷者的自評程度更明顯集中於3分。  
    
    
  
* 網路沉迷尋求協助：  
  針對266位自認有網路沉迷狀況者，只有4.5%「想過」要尋求幫助。  
  

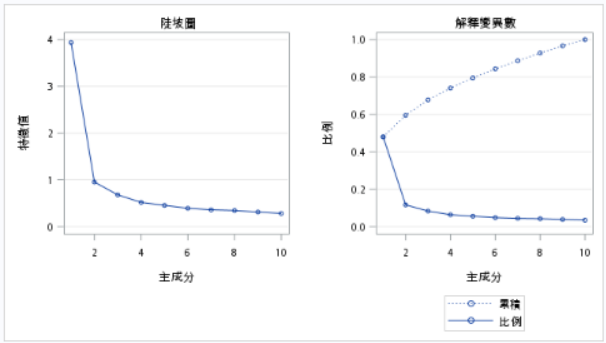
1. 心理健康危險因子

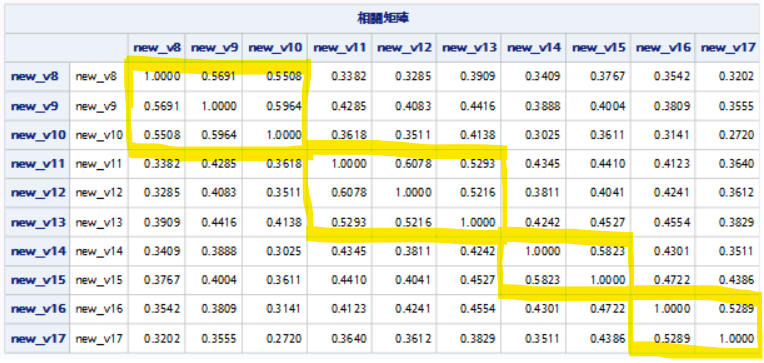
Q28~Q30依照原研究設計分別對應憂鬱、無聊感、課業或工作壓力三項心理健康危險因子，透過5分量表詢問受訪者的身心狀況，1~5分由低至高依序為非常不同意至非常同意。排除遺漏值與不知道/拒答的回答結果後可見，多數受訪者沒有搞不懂工作或課業上的壓力，但有較多人對生活具有無聊感，平均分數相對較高，且第三四分位數來到3分。  
  
憂鬱、無聊感、課業/工作壓力三項分別正面表達困擾的比例分別為7.7%、17.5%、5.0%，又以無聊感回答3分的比例最高。  
  


1. **主成分分析**

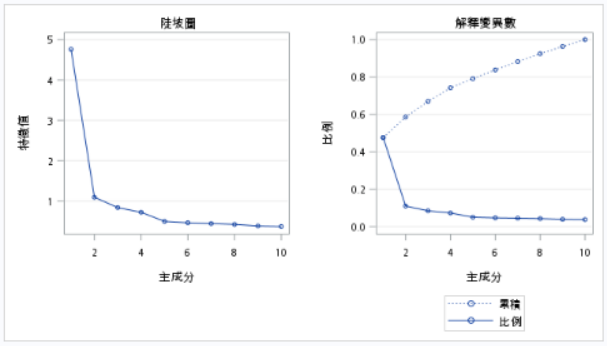
首先針對「網路使用習慣量表」進行主成分分析，看看題組中關鍵的判別變數為何，哪一個維度的指標最能判斷網路使用者是否有網路沉迷的問題。進行主成分分析時，將量表所有子題皆未回答/拒答或遺漏值的受訪者排除，故僅針對1521筆資料進行分析。

首先以均值修正（mean-corrected）資料進行分析，共變異數矩陣中，變數間的互變異數皆小於1，而第一主成分的特徵值為3.936，占總變異量的47.92%，直到第四主成分累積解釋變異量才突破七成，占74.04%。而從特徵像量也可以看到第一主成分與所有子題指標的分數皆維持0.2~0.4之間，並未有特別突出的類別。

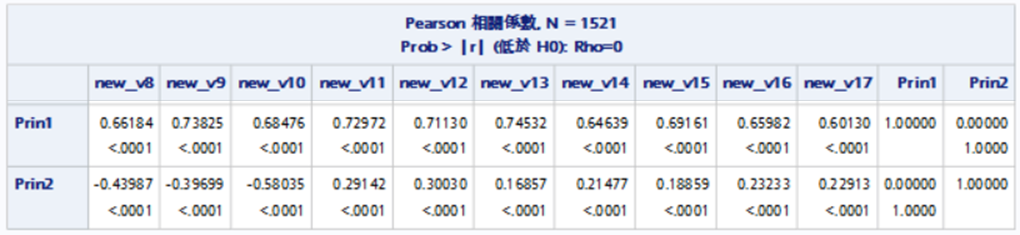
  
   


若以標準化資料進行主成分分析，透過相關矩陣可以看到，變數自己的相關係數固定為1，相同測量目標的題組間相關性更為明顯，彼此之間中度相關（0.5~0.6之間），高於均值修正資料的矩陣值（0.3~0.6之間）。而從特徵值結果可以看到第一主成分可解釋變異量接近過半，47.59%，累積到第四主成分時可解釋變異量為74.17%，並有2個主成分的特徵值大於1，搭配陡坡圖，肘點落在主成分2的位置，**與均值修正資料結果相同，故應保留兩個主成分**。  




從相關係數結果





1. **因素分析**
2. **集群分析**

1. **建立模型準備**

這裡首先我們要先決定Y，其中我們比較關心的議題有V3\_min ( 工作日/上課日使用電子設備時間 )、V4\_min ( 非工作日/上課日使用電子設備時間 )，其中我們有考慮過分開建模型或是平均來做為我的Y。因此這裡先對它們最相關性檢定。

1. 皮爾森相關性檢定

這裡針對相關性的虛無假設為：

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

由此結果可以看出來相關係數 (r)為0.66297，表示存在中等偏高的正相關。而且也可以看到 P 值 < 0.0001 拒絕虛無假設，表示他們彼此之間有顯著的關聯，並且這裡我們是主要想探討整體趨勢，因此後續決定把他們平均來當Y更加合適。

1. Shapiro-Wilk 單變量

這裡定義我們的Y來做為後來預測的變數。

Shapiro-Wilk 單變量檢定這裡我們可以建立虛無假設：

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

由Shapiro-Wilk 常態性檢定<0.0001拒絕虛無假設，此數據不為常態分配。也可以看出偏態為正，表示數據呈現 右偏分佈（右尾較長）。峰態值大於 0，表示數據有較高的集中度，但同時也可能存在較多的極值，後續可以對變數做轉換，但轉換後的數據可能會失去部分解釋意義，因此這裡直接進入羅吉斯回歸分析，避免過多的數據處理對結果的影響。

1. **羅吉斯回歸**
2. 定義Y

這裡我們會把Y分為一天內扣除睡覺時間( 9小時 ) 五分之一以上時間(3\*60)都在使用電子產品，以及五分之一以下的時間在使用電子產品。

一張含有 Rectangle, 螢幕擷取畫面 的圖片

自動產生的描述

由上面兩個的類別可以看的出來數量差不多，接下來就可以進入到羅吉斯回歸的部分。

1. stepwise 挑選變數

一張含有 文字, 數字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

這裡我們挑選前五的變數來做為後續建立模型的主要變數。這些變數的 P-值均小於 0.0002，表明它們對目標變數的影響顯著。

1. 羅吉斯回歸

使用上面的主要變數來進入羅吉斯回歸中。

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述 一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

此時的AIC=1749.455，SC=1754.599，表示模型適配性尚可，適合進行解釋和預測。P 值 < 0.0001，表明至少有部分解釋變數對目標變數有顯著影響。和諧百分比為70.0%，表明模型在預測方面具有良好的能力。c=0.723：表明模型具有一定的區分能力，接近於優秀模型的水準（c≥0.7 通常視為較好的分類能力）。

一張含有 行, 圖表, 繪圖, 文字 的圖片

自動產生的描述

判定模型性能的方法是 ROC 曲線下的面積（AUC）。面積越大，表示模型的預測能力越好。從圖中可以看出，該模型的 AUC 值為 0.7228，曲線下的面積較大，表明該模型具備良好的預測能力。

1. **典型相關**

**第四章 結論與建議**

**第一節　結論**

**第二節　建議**

**第五章 附錄**

**第一節　問卷**

**第二節　問卷編碼**

1. 網路使用習慣量表是基於26 題陳氏網路成癮量表（Chen Internet Addiction Scale，簡稱CIAS）的10 題版本（此處簡稱網路使用習慣量表為CIAS-10）；根據國發會106 年的網路沉迷研究調查結果顯示，26 題的陳氏網路成癮量表與10 題的網路使用習慣量表（CIAS-10）的篩選結論一致。 [↑](#footnote-ref-1)
2. 利用上網率與各縣市的母體分布，估算各縣市預計完成樣本數（上網者+非上網者）。 [↑](#footnote-ref-2)
3. 研究共撥出47,056支電話，接觸到2,900 位合格受訪者，1,975/2,900=68.1%。 [↑](#footnote-ref-3)