

量化資料分析

Chapter 13

Introduction to Social Research 5E

研究方法 二版 基礎理論與技巧

Earl Babbie 原著 林佳瑩 審閱
蔡毓智、邱泯科、陳佳穎、姜馨彥 合譯



13.1 引言

- 量化分析 (quantitative analysis)
 - 為了對觀察所反映的現象進行描述與解釋的目的，而將觀察加以操弄及進行數值化的呈現。
- 大部分的資料一開始都是質化的：因此必須加以數量化，才能進行統計分析。



13.2 資料的量化

- 現今，研究幾乎都是使用電腦軟體（像是 SPSS 或 Micro Case 等等）以進行量化分析。
- 研究者選擇某一套分類標準以過錄資料時，必須考量是否合適於這個研究本身所檢視的理論概念。
- 在原始資料過錄階段，盡可能詳盡記錄資訊細目。



13.2 資料的量化

表 13-1 學生回答項目中涉及「經濟考量者」

	財務上的考量
學費太高	X
停車位不足	
教授們不知道自己在做什麼	
永遠找不到指導教授	
開設課程不足	
宿舍裡太多蟑螂	
要求條件太多	
餐廳食物不潔	
書籍費太高	X
獎助學金不足	X



13.2 資料的量化

表 13-2 學生回答項目中涉及「學術性者」與「非學術性者」

	學術性者	非學術性者
學費太高		X
停車位不足		X
教授們不知道自己在做什麼	X	
永遠找不到指導教授	X	
開設課程不足	X	
宿舍裡太多蟑螂		X
要求條件太多	X	
餐廳食物不潔		X
書籍費太高		
獎助學金不足		X



13.2 資料的量化

表 13-3 涉及「非學術性者」與「行政事物相關」與「校園設施相關者」分

	學術性者	行政事物相關者	校園設施相關者
學費太高		X	
停車位不足			X
教授們不知道自己在做什麼	X		
永遠找不到指導教授			
開設課程不足	X		
宿舍裡太多蟑螂			X
要求條件太多	X		
餐廳食物不潔			X
書籍費太高	X		
獎助學金不足		X	



13.2 資料的量化

□ 13.2.2 建構過錄簿 (codebook)

- 資料處理及分析過程中使用的文件，用以告知資料檔中不同資料題目的位置。
- 通常，過錄簿會指示資料題目的位置及用以代表變項不同屬性代碼的意義。



13.2 資料的量化

POLVIEWS	ATTEND
<p>這些日子以來，我們聽到許多有關「自由取向」與「保守取向」的談論。現在，我將給你看一個 7 點量表 (a seven-point scale)，對於人們所抱持的政治觀點從極端自由取向的——分數 1 ——排列到極端保守取向的——分數 7，在此量表中，你會將自己的政治觀點放置於何處？</p> <ol style="list-style-type: none">1. 極端自由取向的2. 自由取向的3. 稍微自由取向的4. 適中的5. 稍微保守取向的6. 保守取向的7. 極端保守取向的8. 不知道	<p>你上教堂的頻率是</p> <ol style="list-style-type: none">0. 從未參加1. 一年少於 1 次2. 一年 · 或 2 次3. 一年數次4. 每個月一次5. 每個月 2-3 次6. 幾乎每週一次7. 每週一次8. 每週數次9. 不知道、沒有回答

圖 13-1 過錄簿的部分節錄

13.2 資料的量化

□ 13.2.3 資料輸入

- 研究者若想進行量化資料分析，除了必須將資料轉換為數量化格式，還需要把資料轉變為機器可以判讀的格式 (machine-readable format)，才能運用電腦讀取與探尋資料檔。
- 資料輸入的工作，也可能與資料蒐集階段同步完成。
- 最後一旦資料可以完全數量化、並且輸入電腦，研究者就可以開始進行資料分析。

13.3 單變項分析

□ 單變項分析 (univariate analysis)

□ 指的是對單一變項的分析，目的在進行描述。

□ 13.3.1 分配

□ 次數分配 (frequency distribution)

□ 對樣本中被觀察到某個變項的不同屬性出現次數的數字描述。



13.3 單變項分析

SDA Frequencies/Crosstabulation Program
Help: [General](#) / [Recoding Variables](#)

REQUIRED Variable names to specify
Row:

OPTIONAL Variable names to specify
Column:
Control:

Selection Filter(s): Example: age(18-50)

Weight:

TABLE OPTIONS	CHART OPTIONS
<p>Percentaging: <input checked="" type="checkbox"/> Column <input type="checkbox"/> Row <input type="checkbox"/> Total with <input type="text" value="1"/> decimal(s)</p> <p><input type="checkbox"/> Confidence intervals Level: <input type="text" value="95 percent"/></p> <p><input type="checkbox"/> Standard error of each percent</p> <p>Sample Design: <input checked="" type="radio"/> Complex <input type="radio"/> SRS</p> <p><input type="checkbox"/> Statistics with <input type="text" value="2"/> decimal(s)</p> <p><input type="checkbox"/> Question text <input type="checkbox"/> Suppress table <input checked="" type="checkbox"/> Color coding <input type="checkbox"/> Show Z-statistic <input type="checkbox"/> Include missing-data values</p>	<p>Type of chart: <input type="text" value="Bar Chart"/></p> <p>Bar chart options: Orientation: <input checked="" type="radio"/> Vertical <input type="radio"/> Horizontal Visual Effects: <input type="radio"/> 2-D <input checked="" type="radio"/> 3-D</p> <p>Show Percents: <input checked="" type="checkbox"/> Yes</p> <p>Palette: <input checked="" type="radio"/> Color <input type="radio"/> Grayscale</p> <p>Size - width: <input type="text" value="600"/> height: <input type="text" value="400"/></p>

圖 13-2 上教堂的單變項分析



13.3 單變項分析

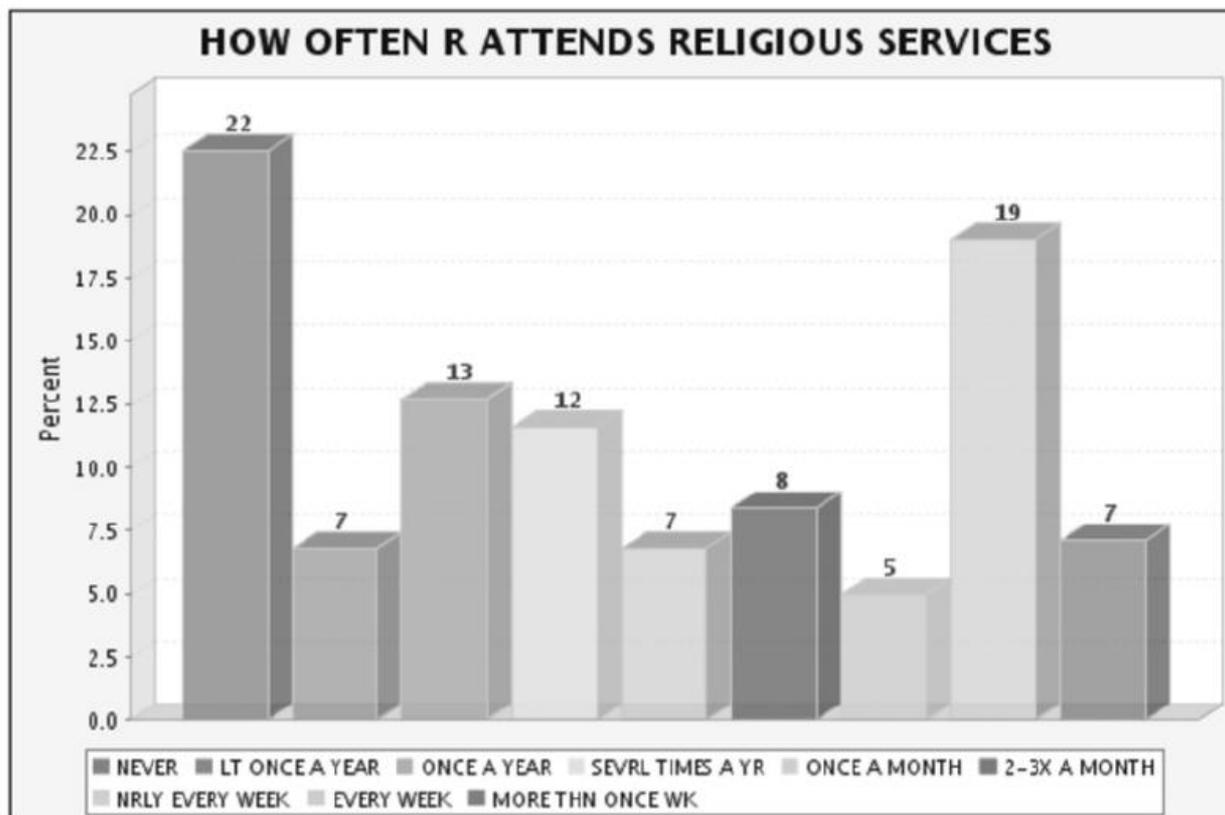


圖 13-3 GSS 資料中上教堂的長條圖



13.3 單變項分析

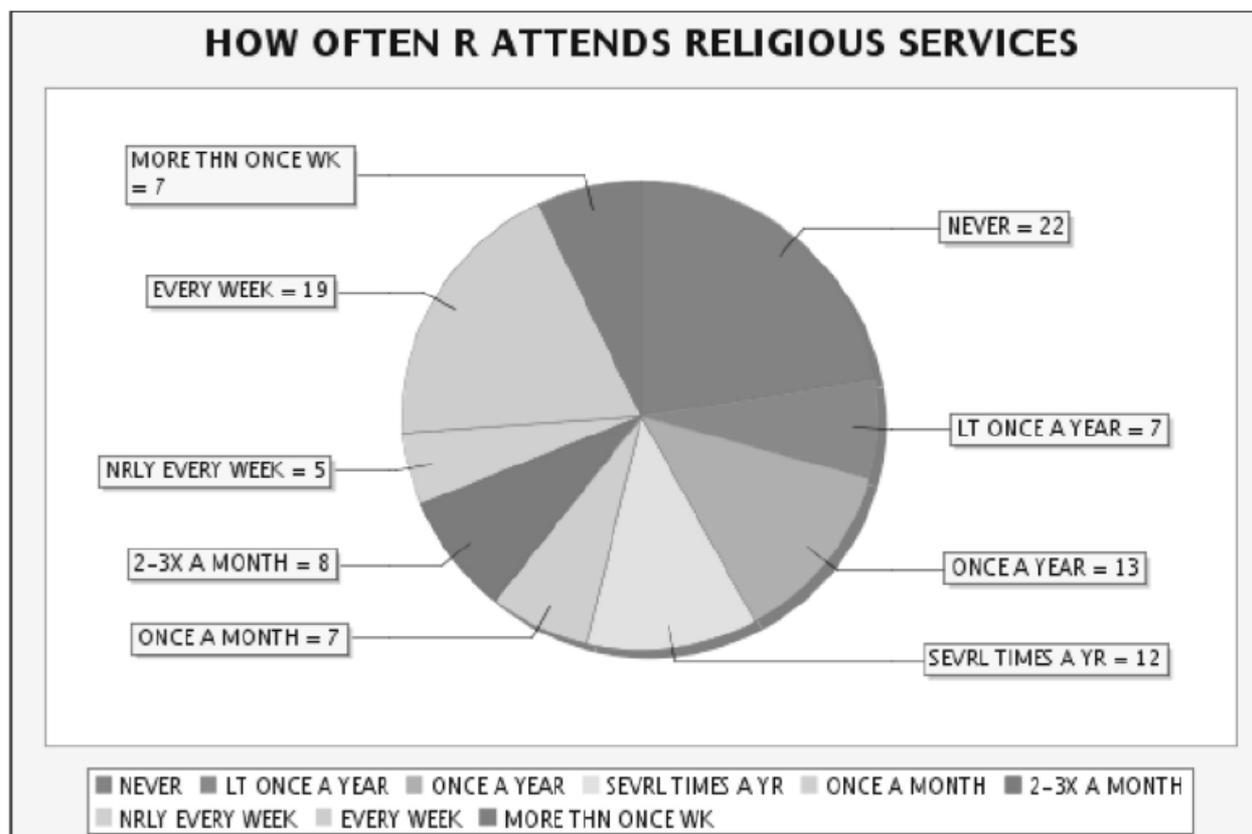


圖 13-4 2006 年 GSS 資料中上教堂的圓餅圖

13.3 單變項分析

表 13-4 2006 年上教堂作禮拜的情形

受訪者上教堂的頻率

數值標示	數值	次數	百分比
從來沒有	0	1,009	22.5
一年少於 1 次	1	305	6.8
每年一次	2	571	12.7
每年數次	3	522	11.6
每月一次	4	307	6.8
每月 2-3 次	5	378	8.4
幾乎每週一次	6	224	5.0
每週一次	7	856	19.0
每週數次	8	321	7.1
	總計	4,492	100.0



13.3 單變項分析

□ 13.3.2 集中趨勢 (central tendency)

□ 平均數 (mean) :

- 由觀察個案數值的總和除上觀察個案的總數所計算而得的平均。

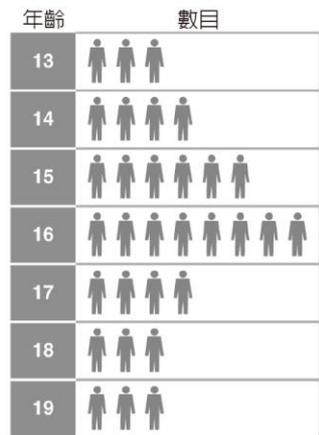
□ 眾數 (mode) :

- 某個最經常出現的屬性用作平均性的代表。

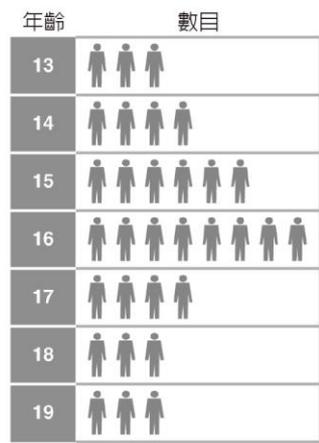
□ 中位數 (Median) :

- 在一個經過次序排序之後的觀察樣本中，位置「中間」位置者被用來作為平均代表。





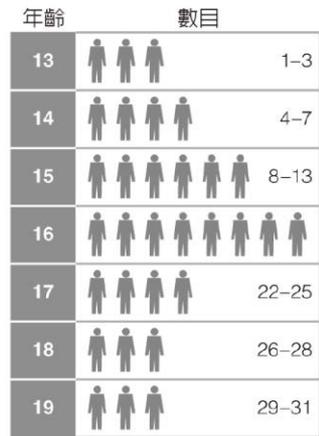
眾數 = 16
 最多的次數



$13 \times 3 = 39$
 $14 \times 4 = 56$
 $15 \times 6 = 90$
 $16 \times 8 = 128$
 $17 \times 4 = 68$
 $18 \times 3 = 54$
 $19 \times 3 = 57$
 492 總計

平均數 = 16.37
 算術平均數

$$\frac{492 \text{ 總計}}{31 \text{ 個案}} = 15.87 + 0.50 = 16.37$$



中位數 = 16.31
 中點

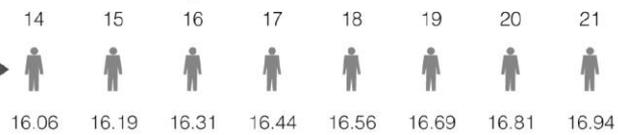


圖 13-5 三種不同的「平均情形」

13.3 單變項分析

□ 13.3.3 離散度 (dispersion)

→ 數值沿著某個中心數值平均分佈的程度。

□ 全距 (range)

□ 最大數值與最小數值的距離。

□ 標準差 (standard deviation)

□ 一個測量平均數周圍離散程度的測量值。

□ 四分位差 (interquartile range)

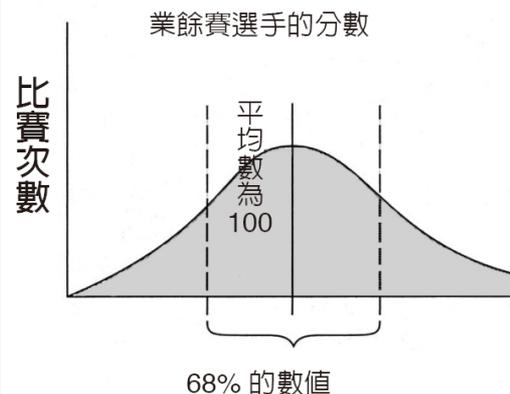
□ 反映中間半數人的分數分布狀況。



13.2 資料的量化



A. 高標準差 = 數值分散



B. 低標準差 = 數值較為集中

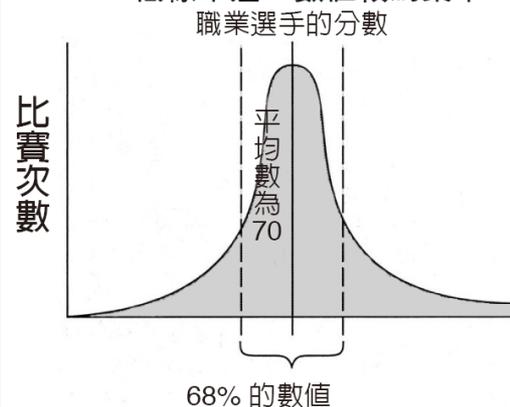


圖 13-6 高與低標準差



13.3 單變項分析

□ 13.3.4 連續與間斷變項

□ 連續變項 (continuous variable)

□ 一個變項的屬性構成了一個穩定的數列。

□ 間斷變項 (discrete variable)

□ 一個變項的屬性彼此之間是相互分離的，或是不連續的。

□ 13.3.5 細緻相對於操控性

□ 盡可能使用超過一種型式以上的資料，藉以獲得更多關於研究資料的訊息。



13.4 分組比較

13.4.1 「合併」反應類目

適切合併反應類目，就可能避免某些危險。

表 13-7 對聯合國的態度：「聯合國解決其所面臨的問題的表現如何？」

	西德	英國	法國	日本	美國
很好	2%	7%	2%	1%	5%
好	46	39	45	11	46
不好	21	28	22	43	27
非常不好	6	9	3	5	13
不知道	26	17	28	41	10

資料來源：「5-Nation Survey Finds Hope for U.N.,」 *New York Times*, June 26, 1985, p. 6.

表 13-8 化約極端值後的情形

	西德	英國	法國	日本	美國
好或很好	48%	46%	47%	12%	51%
不好或非常不好	27	37	25	48	40
不知道	26	17	28	41	10



13.4 分組比較

□ 13.4.2 處理「不知道」

- 處理不知道選項時，一個簡單的方法，就是把回答「不知道」的受訪者排除，然後重新計算百分比。

表 13-9 省略「不知道」者

	西德	英國	法國	日本	美國
好或很好	65%	55%	65%	20%	57%
不好或非常不好	35	45	35	81	44



13.4 分組比較

□ 13.4.3 質化研究中的數量化描述

- 進行社會研究時，結合不同的研究途徑，將會極有幫助且更為有力，尤其結合質化研究與量化研究最具潛力與效能。



13.5 兩變項分析

□ 兩變項分析 (bivariate analysis)

- 同時對兩個變項進行的分析，用以決定二者之間的實證關係。
- 百分比表的建立或是簡單相關係數是都是兩變項分析的例子。

表 13-10 2006 年由男性與女性所報告的宗教參與情形

	男性	女性
每週	26%	35%
較少	74	65
100% =	(2,049)	(2,443)

Source: General Social Survey, 2006, National Opinion Research Center.

本頁以任何形式轉載、重製或編及列印。



13.5 兩變項分析

□ 13.5.1 建立交叉表的百分比分配

- 1. 當交叉表以直行方向計算百分比時，跨越橫列比較與閱讀百分比數據。
- 2. 相反地，若交叉表示從橫列方向計算百分比，則跨越直行閱讀與比較百分比數據。

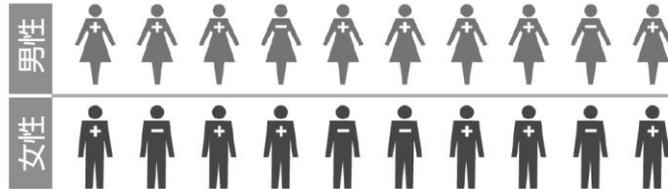




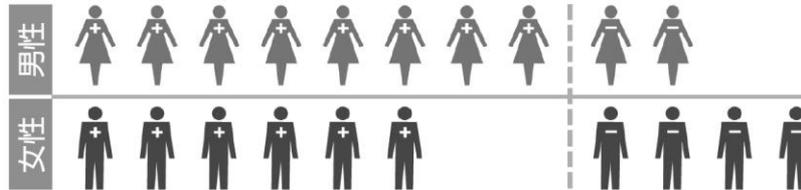
a. 若干的男性與女性個案，其中，有些人支持(=)、另一些反對(≠)兩性平等



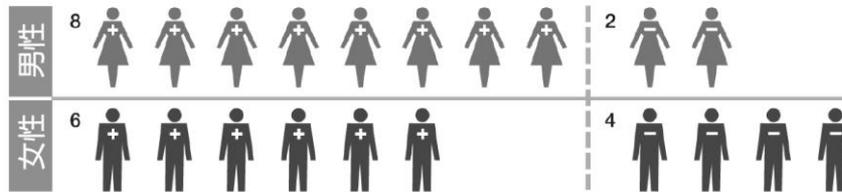
b. 將男性與女性(自變項)分開。



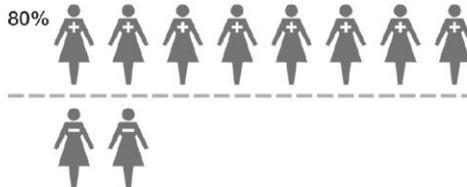
c. 在每一個性別次群體內，再將哪些支持及反對者(應變項)分開。



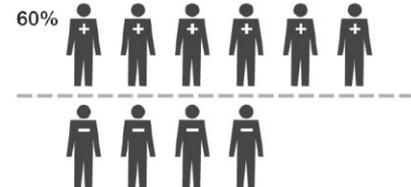
d. 在表內每一方格計算人數。



e. 計算女性中支持者的百分比



f. 計算男性中支持者的百分比



g. 結論 雖然不論男性或女性個案，大多數是支持兩性平等，女性較之男性更為可能支持。所以，性別就是造成對於兩性平等態度差異的原因之一。

支持
反對
合計

女性	男性
80%	60%
20	40
100%	100%

圖 13-7 計算百分比 (續)



13.5 兩變項分析

表 13-11 對於大麻合法化議題之報紙編輯方針（假設的資料）

對於大麻合法化 議題之編輯方針	社區規模 100000 人以上	
	100000 人以下	100000 人以上
贊成	11%	32%
中立	29	40
反對	60	28
100%=	(127)	(438)



13.5 兩變項分析

13.5.2 建構與閱讀兩變項交叉表

- 建構解釋性之兩變項交叉表的步驟如下：
 - 1. 根據自變項的屬性，將研究對象分組。
 - 2. 在每一個次群體組別內，分別描述依變項的屬性。
 - 3. 最後，根據被給定的依變項的屬性，比較自變項內不同組別的差異。
- 列聯表
 - 以百分比方式呈現變項之間的關係，通常被用來顯現自變項對依變項產生的效果。



13.5 兩變項分析

表 13-12 2006 年性別、年齡與宗教參與的多變項關係

「請問您多久參加宗教活動？」

	40 歲以下		40 歲及以上	
	男性	女性	男性	女性
一個星期左右*	24%	32%	33%	48%
較不常	76	68	67	52
100%=	(325)	(383)	(323)	(452)

*「一個星期左右」包括：「每週一次以上」、「每週」、和「幾乎每週」。

資料來源：從 1973、1983、及 1993 年之 GSS 抽取隨機樣本計算。



13.6 多變項分析簡介

□ 多變項分析 (multivariate analysis)

- 同時分析數個以上的變項之間關係的分析。
 - 多變項分析納入考量一個以上的自變項；所以不再只是以一個自變項而解釋依變項的變異。
 - 反之我們尋求從更多的自變項來解釋依變項。

表 13-13 表 13-12 的簡化表

	每週上教堂禮拜的百分比	
	男性	女性
40 歲以下	24 (325)	32 (383)
40 歲及以上	33 (323)	48 (452)



13.7 社會診斷

- 社會診斷是一種量化的分析技巧，用以了解某些諸如族群或性別歧視等社會問題的性質。

表 13-14 1984 年性別、工作情形與收入 (21-64 歲的全職工作者)

任職目前工作的年數	平均時薪		女性／男性比率
	男性	女性	
少於 2 年	\$ 8.46	\$6.03	.71
2-4 年	\$ 9.38	\$6.78	.72
5-9 年	\$10.42	\$7.56	.73
10 年及以上	\$12.38	\$7.91	.64

資料來源：Bureau of the Census, Current Population Reports, Series P-70, No. 10, *Male-Female Differences in Work Experience, Occupation, and Earning, 1984* (Washington, DC: U.S. Government Printing Office, 1987): 4.



13.8 量化資料分析與倫理

- 無偏誤分析與進行報告在量化研究中與在質化研究中相同，都是重要倫理考量。
- 在量化資料分析與進行報告時，必須保護被研究者的隱私。





The End

