

編碼：RES-98-04

行政院主計處委託研究

工商普查抽樣方法效能之研究

受委託單位：中華市場研究協會

行政院主計處編印

印製日期：98年12月

目錄

目錄.....	一
摘要.....	三
表目錄.....	五
圖目錄.....	七
第一章 緒論.....	1
第一節 研究背景與動機.....	1
第二節 研究目的.....	2
第三節 研究流程.....	3
第二章 文獻探討.....	4
第一節 普查應用名詞定義.....	4
第二節 我國工商及服務業普查抽樣設計發展.....	8
壹、85年普查抽樣設計.....	8
貳、90年普查抽樣設計.....	8
參、95年普查抽樣設計.....	9
第三節 各國普查辦理情形.....	13
第四節 各國經濟調查之抽樣及分層方式.....	19
第五節 各種分層方式比較.....	26
第三章 研究方法.....	30
第四章 資料分析.....	40
第一節 母體資料分析.....	40
壹、95年普查母體資料.....	40
貳、歷次普查母體結構分析.....	41
第二節 樣本結構分析.....	43
壹、95年普查樣本資料.....	43
貳、列聯表分析.....	45
參、關聯性分析.....	46
第三節 95年普查母體與樣本資料分析.....	53
壹、抽出率.....	53
貳、層數分析.....	54

第四節 樣本精確度分析.....	56
壹、營業收入.....	56
貳、從業員工人數.....	57
參、資產淨額.....	57
第五節 行業別的分類（以製造業為例）.....	59
第六節 交叉表中變項組合之次數分析.....	63
第七節 交叉表中變項類別之特定對應組合關聯分析.....	68
第八節 結果與討論.....	74
第五章 結論與建議.....	78
第一節 結論.....	78
第二節 建議.....	80
參考文獻（中文部分）.....	82
參考文獻（英文部分）.....	83
附錄 1 連續營運特徵值間的相關係數—印刷業.....	85
附錄 2 連續營運特徵值間的相關係數—服裝及其配件批發業.....	87
附錄 3 印刷業營業收入分布—樣本資料.....	89
附錄 4 印刷業營業收入分布—樣本資料取 0.134912 次方轉換.....	90
附錄 5 印刷業營業收入分布—母體資料.....	91
附錄 6 印刷業營業收入分布—母體資料取 0.134912 次方轉換.....	92
附錄 7 CART 樹狀圖.....	93
附錄 8 CART 樹狀圖—從業員工薪資.....	94
附錄 9 地區、營業收入及行業列聯表(4×5×6).....	95
附錄 10 地區、營業收入及行業樣本資料對數線性模型參數估計 ...	96

摘要

95 年普查首次爰採用多變量自動互動檢視法(Chi-square Automatic Interaction Detection, CHAID)及即時線上分析處理(Online Analytical Processing, OLAP Cubes)進行分層設計，抽樣家數較 90 年普查大幅減少，除分工單位外，地方普查單位抽樣家數降至 4 萬 8 千家，對於調查品質控管及成本效益之提升，均有正面效益。本研究除先就歷年普查問項、相關定義及抽樣方法內涵略予介紹外，尚蒐集世界各國工商業調查及運作模式等資訊，期藉對照比較，從中汲取優點；並為擷節調查成本，維持統計代表性及確度。鑒於該抽樣設計係國內企業面調查首創，允有必要於普查作業實施後，運用相關資料進行深入評估研究，俾確實了解其成效，精進抽樣技術，並作為下次普查規劃基礎。

藉母體與樣本分布結構，以及運用列聯表分析、關聯分析、對數線性模式、對應分析與分類及迴歸樹分析(Classification and Regression Tree, CART)等統計方法，探究樣本重要特徵值之特性，俾驗證 95 年抽樣分層作業之效能。

近年來產業變遷迅速，企業獲利能力未必與投入存在絕對之比例關係，經實證分析，本次普查運用多變數分層概念，確能有效降低層內變異，且選取變數亦具備足夠之妥適性，惟仍發現有樣本精確度及空層問題，爰就分析結果針對分層、樣本配置過程及樣本規模制定等問題予以檢討，並引申本文之結論及未來改進方針，俾臻研究初始目的。

表目錄

表 2.2	95 年普查分層方式.....	11
表 2.3.1	美國經濟普查行業別問卷類型及合計數.....	14
表 2.3.2	美國普查局工商母體資料庫格式內容.....	16
表 2.4	截略點選取準則.....	20
表 2.5	集群分析.....	28
表 3.1	第 i 列剖面權值.....	38
表 3.2	第 j 行剖面權值.....	38
表 4.1.1	各大類行業母體資料分布.....	41
表 4.1.2	75 年至 95 年普查母體資料CHAID分析.....	42
表 4.2.1	各大類行業樣本資料分布.....	44
表 4.2.2	大行業與地區別列聯表.....	45
表 4.2.3.1	印刷業營業收入與各連續特徵值之相關係數.....	47
表 4.2.3.2	服裝及其配件批發業營業收入與各連續特徵值相關係數.....	48
表 4.2.3.3	類別變數定義.....	49
表 4.2.3.4	印刷業以縣市別預測各層類別變數之相關情形.....	49
表 4.2.3.5	服裝及其配件批發業以縣市別預測各層類別變數之相關情形.....	50
表 4.2.3.6	印刷業營業收入與各類別變數之相關程度.....	51
表 4.2.3.7	服裝及其配件批發業營業收入與各類別變數之相關程度.....	52
表 4.3.1	各大類行業抽出率分布.....	54
表 4.3.2	各大類行業層數分布.....	55
表 4.4.1	母體與樣本平均每企業營業收入.....	56
表 4.4.2	母體與樣本平均每企業從業員工人數.....	57
表 4.4.3	母體與樣本平均每企業資產淨額.....	58
表 4.5.1	CART分析變數屬性說明—製造業.....	60
表 4.5.2	CART分析變數初步選取—製造業.....	62
表 4.5.3	I_{x_i} 及 $R^2 x_i$ 值—製造業.....	62
表 4.6.1	95 年普查抽樣資料營業收入(y)DH分層.....	63
表 4.6.2	最大概似變異數分析—飽和模型.....	64
表 4.6.3	最大概似變異數分析—非飽和模型(移除三個效應之交互作用效應項).....	65

表 4.6.4	最大概似變異數分析—非飽和模型(移除所有交互作用效應)	65
表 4.7.1	列聯表—固定製造業、礦業及土石採取業	68
表 4.7.2	摘要表—固定製造業、礦業及土石採取業	69
表 4.7.3	列點數綜覽—固定製造業、礦業及土石採取業	69
表 4.7.4	行點數綜覽—固定製造業、礦業及土石採取業	69
表 4.7.5	相關測度值—固定製造業、礦業及土石採取業	69
表 4.7.6	摘要表—地區、行業及營業收入	71
表 4.7.7	列點數綜覽—地區、行業及營業收入	72
表 4.7.8	行點數綜覽—地區、行業及營業收入	72
表 4.7.9	相關測度值—行業為製造業、礦業及土石採取業	73

圖目錄

圖 1.3	研究流程.....	3
圖 2.3.1	工商母體資料庫－義大利.....	17
圖 2.3.2	工商母體資料庫及衛星資料庫關聯圖.....	18
圖 4.5.1	製造業樣本資料營業收入分布.....	59
圖 4.5.2	製造業樣本資料 $\ln(\text{營業收入}+1)$ 分布.....	60
圖 4.7.1	地區與營業收入分布圖－固定製造業、礦業及土石採取業	70
圖 4.7.2	地區、行業及營業收入分布圖.....	73

第一章 緒論

第一節 研究背景與動機

工商及服務業普查旨在藉由 5 年 1 次的總體檢，了解全國產業發展特徵與趨勢，係當前產業政策規劃之重要參據。由於普查作業需耗用大量資源，因此其品質、效率與方法，備受各界關注。

隨著工商業多元化發展，經濟情勢變化速度加劇，各界對產業資訊涵蓋範圍之需求日增，惟普查工作亦面臨調查環境惡化及調查成本提高的挑戰，傳統普查須逐家查填數十問項之作業方式實施難度倍增，所幸抽樣技術日益成熟，近 3 次普查爰將全面訪查項目大幅簡化，而原全面訪查之調查項目，亦改以抽樣調查方式執行。85 年普查係將從業員工 3 人以下廠商改以抽樣調查，抽出率 3%；4 人以上廠商以細類行業家數少於 3 千家者全查，大於 3 千家者抽樣調查，抽出率約 7 成，總抽查家數約 25 萬家；90 年普查時，營業收入 1 千萬以上廠商維持全查，未達 1 千萬者抽樣調查，抽出率 1.6%，總抽查家數約 14 萬家，經由此項改革，作業成本及廠商填報負擔均較全面訪查大幅降低。

95 年普查首次採用多變量自動互動檢視法(Chi-square Automatic Interaction Detection, CHAID)及即時線上分析處理(Online Analytical Processing, OLAP Cubes)進行分層設計，抽樣家數較 90 年普查大幅減少，除分工單位外，地方普查單位抽樣家數降至 4 萬 8 千家，對於調查品質控管及執行效率提升，均有正面效益，期運用有限的調查成本，維持統計代表性及統計確度。該抽樣設計係國內企業面調查首創，允有必要於普查作業實施後，運用相關資料進行深入評估研究，俾確實了解其成效，精進抽樣技術，並作為下次普查規劃基礎。

第二節 研究目的

抽樣設計之良窳攸關抽樣調查之品質，抽樣設計須能掌握母體特性及結構，以確保樣本能充分反映母體內涵。政府各機關每年辦理 70 餘項企業面統計調查，為陳示各行業特徵，調查抽樣設計多採分業分層抽樣法，依行業別及主要變數進行分層，以達最佳推估效果。

分層準則的訂定，係分層抽樣設計之關鍵，惟企業經營日趨多元，影響營運成效因素愈趨複雜，以往以單一變數訂定分層準則之方式，愈有難以陳示產業結構變遷內涵之憾。95 年工商及服務業普查抽樣作業爰將分層方法由以營業收入單變數最適分層法改為多變數自動互動檢視法及即時線上分析處理，分層準則中增加從業員工人數、資產淨額及縣市鄉鎮別等變數，期強化層間異質、層內同質之特性，以提升樣本代表性。

95 年普查母體家數達 100 餘萬家，為詳盡陳示各行業經營活動狀況，抽樣設計係依最新行業標準分類近 500 項細類行業為副母體，衡酌經費、人力限制及統計確度，將總樣本數大幅減少至 4 萬餘家，相關作業執行方法、樣本配置及結果陳示等，亦因應抽樣設計改變而配合調整。鑒於抽樣作業複雜且工作量龐鉅，為持續精進普查作業，本研究爰結合普查資料深入檢視評估本次抽樣調查樣本特性及與母體結構之一致性，以了解本次普查抽樣調查樣本代表性及推估效益，並針對現行抽樣方法及各步驟流程詳予評估，提出改進建議，作為下次普查抽樣設計構想研訂之依據。

第三節 研究流程

本文共分五章，除本章外，其餘各章內容及架構如次：

第二章：除普查應用名詞介紹外，並蒐集相關文獻，簡要說明我國普查抽樣設計發展、各國普查辦理情形、抽樣及分層方式，及企業面調查常用抽樣設計、統計方法及分層方式。

第三章：闡述本文應用之研究方法。

第四章：針對 95 年普查資料進行分析，運用前章介紹之各種統計方法，評估抽樣規劃與執行結果預期目標之達成度。

第五章：歸納前章研究結果，引申本文結論，並提出改進方針及建議。

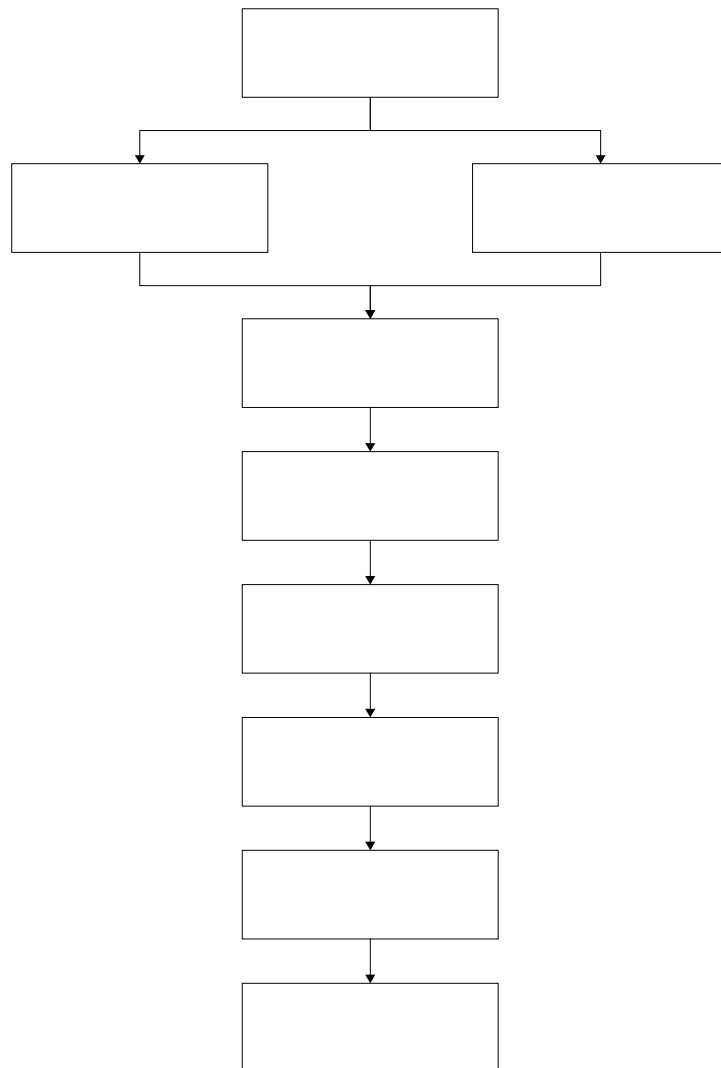


圖 1.3 研究流程

第二章 文獻探討

第一節 普查應用名詞定義

- 一、 **企業單位**：以一個場所或多個場所結合成一個事業單位或獨立的經營體，從事一種或多種經濟活動，自行決定經營方針、資金運用等，備有經營帳簿，並自負盈虧者，視為一個企業單位。
- 二、 **行業歸類**：「行業」係指經濟活動的種類，包括從事生產各種有形物品與提供各種服務的經濟活動。依中華民國行業標準分類第 8 次修訂，其歸類原則如下：
 - (一) **直接從事生產（包括貨品之生產、銷售及服務之提供）之場所**：若同一場所從事數種性質不同的經濟活動（含年中變動、季節不同）時，依序按照下列原則，歸類主要經營業務：
 1. 全年營業或銷售收入總額最多之產品或服務。
 2. 投入從業員工人數較多之產品或服務。
 3. 投入資產設備較多之產品或服務。
 - (二) **未直接從事生產之場所**：未從事生產、銷售或提供服務之企業總管理單位、總管理處、企業分支管理機構，歸入企業總管理機構，否則依主要經濟活動歸類。
 - (三) **95 年內有營業，年底已停、歇業之場所**：如有從業人員可資查詢者，依其停、歇業前之經濟活動歸類。
 - (四) **95 年底籌辦中之場所**：依其預計之經濟活動歸類。
- 三、 **普查對象**：凡 95 年 12 月 31 日普查標準日，於臺閩地區從事工商及服務業普查行業範圍經濟活動之場所單位，其設有固定處所及從業人員者，均為本普查之對象。若於 95 年內曾營業，而在普查標準日以前已停、歇業，或年底尚在籌辦中之場所，如有從業人員可資查詢者，仍應列入本普查對象。根據以上定義，普查對象應具備下列三項要件：

- (一)有營業固定場所。
- (二)備有專用營業設備，並置有從業人員。
- (三)從事本普查行業範圍之經濟活動。

四、**普查範圍**：不包括各企業設置在國外的分支單位，以及從事農林漁牧業等分支單位資料。

- (一)**區域範圍**：以臺閩地區為區域範圍，包括各直轄市、臺灣省各縣(市)(含東沙、南沙群島)及福建省金門縣、連江縣。
- (二)**行業範圍**：依據最新中華民國行業標準分類，凡從事工業及服務業二大部門行業均屬之，惟不包括零售或餐飲攤販、信託、基金及其他金融工具、退休基金、研究發展、公共行政及國防(強制性社會安全除外)、教育服務(補習班除外)、創作業、圖書館、檔案保存、博物館及類似機構、宗教、職業及類似組織與家事服務等服務業。
- (三)**對象範圍**：凡經營前揭範圍之企業及場所單位，不論公營或民營，公司組織或非公司組織，已登記或未登記，亦即符合工商及服務業普查之對象者均屬之。
- (四)**資料時期**：「全年」係指 95 年 1 月 1 日至 12 月 31 日之累計數，「年底」者係指 95 年 12 月 31 日之靜態資料。如「從業員工薪資」係指 95 年全年支付員工之薪資。

五、**問項內容定義**：

- (一)**僱用員工**：指該單位年底尚在僱用之員工人數，包括
 1. 95 年 12 月底支領薪資且在職之僱用員工(含常僱員工及臨時員工)。
 2. 支領固定薪資的資本主及家屬從業者。
 3. 參加作業之契約員工、長期在廠工作之計件工。
 4. 建教合作之工讀生、學徒，外籍勞工等。

5. 從事勞動派遣業務之企業，應包括派遣至其他企業工作之員工。

但不包括：

1. 僅支領車馬費，而未實際參加營運作業之董監事、理事及顧問等人員。
2. 留職停薪、全月未參加工作者。
3. 不在廠地工作之計件工作者（即委託廠外家庭包工人數不予查填）。
4. 使用之派遣勞工。
5. 長期服務於國外其他企業及關係企業人員，係指國內企業僱用之員工長期服務於國外其他企業（含子公司或關係企業），且薪資由國內企業支付之情形，例如長期派駐大陸子公司之管理幹部，其薪資係由臺灣母公司支付或臺灣及大陸兩邊同時支領。

(二) 從業員工：係指年底支領薪資且在職之僱用員工，以及不支領固定薪資之自營作業及無酬家屬工作者。

(三) 從業員工薪資：指支付給 95 年內所有曾僱用之常僱員工及臨時員工，於工作期間內之全部薪資報酬，尚包括支付年底已離職者的薪資，自營作業及無酬家屬工作者的薪資及製造業全年委外家庭包工工資。包括本薪、固定津貼（如房租津貼、水電費、伙食費、交通費、實物折值、膳宿折值等）、加班費、各類獎金（工作獎金、年節獎金、年終獎金、不休假獎金等），以及其他津貼（如誤餐費、差旅津貼、交際津貼）。不包括非薪資報酬（福利津貼），如下列各項：

1. 撫卹金、資遣費。
2. 退休金及提存。
3. 各項保險費：勞工保險費、全民健保費、其他保險費等。
4. 其他福利補助費（如教育補助、結婚及生育補助等）。

- (四) **勞動派遣**：亦稱「人力派遣」，係指從事人力派遣業務之企業，與有用人需求的企業或單位簽訂勞務契約，將其僱用之員工派至用人單位，接受用人單位指揮監督及工作分派，提供勞務，而向用人單位收取勞務費或服務費用。
- (五) **全年各項支出**：全年支出係按權責發生制計算，即指企業為經營業務全年所支付之一切費用，包括應付未付款項，但不包括預付與暫付款項及經營上的資本支出（如購置土地或機械設備、新建工程、大修機械等費用）。包含營業成本、營業費用以及非營業費用及損失。
- (六) **全年各項收入**：係按權責發生制計算，即包括全年各項收入外，尚包括應收未收款項，但不包括預收款項。包含營業收入及非營業收入。
- (七) **研究發展支出**：指為改進生產、銷售或服務技術、開發新產品，而購置固定資產的資本性支出及支付的人事、原材物料、維護費、業務費、旅運費等有關費用性支出。
- (八) **員工訓練支出**：指為辦理員工訓練所購置固定資產等資本性支出與支付員工訓練之講師費、場地費、派訓報名費、訓練部門之人事、業務等有關費用性支出。
- (九) **市場行銷支出**：指廣告、市場研究、包裝設計、交際公關，以及行銷部門之人事及業務等費用性支出與購置行銷部門硬體設備之資本性支出。
- (十) **電腦軟體及資料庫購買支出**：指各部門對於電腦軟體及資料庫之相關支出，包括購入成本及租金費用等。
- (十一) **專業技術交易金額**：指經由技術合作、技術授權等方式，購入或銷售技術之金額，包括專利（採購、銷售）、專利授權、專門技術（非專利）、模型和設計、商標（包括經銷權）及技術服務等。

第二節 我國工商及服務業普查抽樣設計發展

我國工商及服務業普查自 43 年創辦迄今，已屆 11 次。為因應國內外經濟情勢變化，調查內容不斷充實，調查設計方法亦配合持續精進。85 年時由於調查成本等因素影響，規劃設計出現重大變革，原本逐家查記普查完整問項方式改以抽樣方式進行，因此，普查結果資料品質除須考量調查執行因素外，尚有抽樣誤差應予控制，抽樣設計作業爰成為普查作業規劃重點工作之一。

壹、85 年普查抽樣設計

一、 抽樣母體：以本處工商業母體資料檔作為抽樣母體。

二、 抽樣方法：

(一) 普查對象從業員工 4 人以上者全查。

(二) 普查對象從業員工 3 人以下者採分業（細類行業）分層隨機抽樣法，以從業員工為分層變數，分為 1 人、2 人及 3 人等三層，各層抽出率以 1% 為原則，如該業家數少於 20 家者，則予全查。

(三) 若細類行業家數大於 3000 家者，酌減填表家數。

(四) 未填普查表單位於普查實地判定階段，查報使用土地及建築物面積、員工人數、全年收入及支出、實收資本額組距資料，作為抽樣推估基礎。

三、 樣本家數：

(一) 從業員工 4 人以上者計 30 萬家，實際抽查 21 萬家。

(二) 從業員工 3 人以下者，實際抽出率 3.5%，計抽出 2 萬 6 千家。

(三) 加計另為產業關聯表所需應用之抽樣長表樣本 2 萬家，合計約 25 萬 6 千家。

貳、90 年普查抽樣設計

一、 抽樣母體：以本處工商業母體資料檔作為抽樣母體。

二、 抽樣方法：

- (一) 普查對象營業收入 1 千萬以上者全查。
- (二) 普查對象營業收入 1 千萬以下者採分業（細類行業）分層隨機抽樣法，以營業收入為分層變數，分為 1 百萬以下、1 百萬至 5 百萬及 5 百萬以上等三層，各層抽出率以 1% 為原則，如該業家數少於 20 家者，則予全查。
- (三) 若細類行業家數大於 3000 家者，酌減填普查表（完整普查問項）家數。
- (四) 未填普查表單位於普查實地訪查階段，填報普查簡易表，包括一般概況、使用電腦情形、使用土地及建築物面積、員工人數、全年收入及支出資料，作為抽樣推估基礎。

三、 樣本家數：

- (一) 營業收入 1 千萬以上者全查，計 10 萬 5 千家。
- (二) 營業收入 1 千萬以下者，實際抽出率 1.6%，計抽出 1 萬 5 千家。
- (三) 加計另為產業關聯表所需應用之抽樣長表樣本 2 萬家，合計 14 萬家。

參、95 年普查抽樣設計

- 一、 抽樣母體：以本處工商業母體資料檔底冊，連結相關公務及統計調查檔案，更新檢核後，以最新之工商業母體資料檔作為抽樣母體。

二、 抽樣方法及樣本家數：

- (一) 細業別家數小於 10 家者，予以全查。
- (二) 醫院、診所、經濟部工業產銷存調查對象、金融業分工調查企業，亦予以全查。

(三) 全查層：以營業收入此項變數，採 M.A.Hidiroglou 法計算截略點(變異係數 $cv=0.05$)，截略點以上全查。

(四) 扣除全查層樣本後，以各細類行業為副母體，在 95%信心水準下，決定各細類行業抽樣家數。

第 i 細業需用樣本數：

$$n_i = \frac{Z^2 \left(\sum_{j=1}^L N_{ij} S_{ij} \right)^2}{\varepsilon^2 (N_i \times \overline{X}_i)^2 + Z^2 \times \sum_{j=1}^L N_{ij} S_{ij}^2}$$

符號說明：

N_i : 第 i 細業母體家數。

N_{ij} : 第 i 細業第 j 層母體家數。

\overline{X}_i : 第 i 細業母體平均數。

S_{ij} : 第 i 細業第 j 層母體標準差。

L : 層數。

Z : $\alpha=0.05$ 之常態標準值。

ε : 相對誤差率，設為 0.05。

(五) 各細業抽樣家數及全查樣本家數合計即為本次抽樣調查家數，原則上全查樣本及經濟部工業產銷存調查對象查填抽樣調查乙表，餘則填報抽樣調查甲表，惟為減輕廠商填報負荷，仍視各大業填報抽樣調查乙表之家數調整之。

(六) 抽樣方式採分業(細類行業)分層系統抽樣法(Systematic Sampling)，按各細類行業家數，分採 OLAP Cubes 或 CHAID 分層。

表 2.2 95 年普查分層方式

	1000 家以下細類行業	1000 家以上細類行業
分層方法	OLAP Cubes 分層法	CHAID 分層法
分層變數	營業收入、從業員工人數	準則變數：營業收入 特徵變數：從業員工人數、資產淨額、縣市鄉鎮
層數之決定	以 D-H 決定分層點後，直接交叉分層	每一層至少應有 3~5 家抽查樣本，控制層數，最多分 10 層。
各層樣本配置	Neyman's allocation 1. 第 i 業第 j 層樣本數 $n_{ij} = n_i \times \frac{N_{ij}S_{ij}}{\sum_{j=1}^L (N_{ij}S_{ij})}$ 2. 第 i 業第 j 層樣本抽出率： $f_{ij} = \frac{n_{ij}}{N_{ij}}$	Neyman's allocation 1. 第 i 業第 j 層樣本數 $n_{ij} = n_i \times \frac{N_{ij}S_{ij}}{\sum_{j=1}^L (N_{ij}S_{ij})}$ 2. 第 i 業第 j 層樣本抽出率： $f_{ij} = \frac{n_{ij}}{N_{ij}}$

符號說明：

- N_i : 第 i 細業母體家數。
- N_{ij} : 第 i 細業第 j 層母體家數。
- S_{ij} : 第 i 細業第 j 層母體標準差。
- L : 層數。
- n_{ij} : 第 i 細業第 j 層樣本家數。
- n_i : 第 i 細業樣本家數。 $n_i = \sum_{j=1}^L n_{ij}$

(七) 樣本總數約為 7 萬 3 千家(其中縣市調查約 4 萬 8 千家，餘為分工單位調查對象)，各類抽樣表樣本數如次：

1. 製造業(含礦業及土石採取業、電力及燃氣供應業及用水供應及污染整治業) 約 23,000 家。

2. 營造業約 3,100 家。
3. 商業(包含批發及零售業、住宿及餐飲業)約 16,000 家。
4. 運輸及倉儲業約 1,500 家。
5. 金融及保險業 1,200 家。
6. 服務事業 (包含資訊及通訊傳播業，不動產業，專業、科學及技術服務業，支援服務業，強制性社會安全，教育服務業，醫療保健及社會工作服務業，藝術、娛樂及休閒服務業，其他服務業) 約 31,000 家。

第三節 各國普查辦理情形

一、經濟統計資訊取得方式

聯合國統計局於 2006 年針對世界各國經濟統計資訊取得方式進行書面調查，答填者計 114 國(96 國未答填，其中 90 國為開發中國家)，其經濟統計資訊之取得，大致可分為以下兩種方式：

- (一) **每 5 年辦理 1 次經濟普查**：即我國之工商及服務業普查，計 39 國答填此項，其中 31 國為開發中國家，如中國、南韓、印度等；6 國為已開發國家，包括美國、日本、義大利、荷蘭、瑞士及列支敦士登，此 6 國雖因應國家經濟帳(national accounts)之作業需求，同時擁有「工商母體資料庫(Business Register)」及各行業別定期經濟調查之體系，惟穩健(robust)統計過程須仰賴多重獨立資料來源，每年產出之估計值數列方可加以比較，故仍持續辦理經濟普查。

就辦理規模、過程、方法及內容而論，美國普查局辦理之經濟普查，實為本作業模式之翹楚，亦為我國工商及服務業普查之主要參考對象。其早年係採各行業分散辦理普查之方式，自 1954 年起方予整合，迄 2007 年計已辦理 12 次。依據美國 2002 年經濟普查辦理歷史手冊(Procedural History of the 2002 Economic Census)，美國經濟普查約涵蓋 2 千 4 百萬場所對象，調查方式以郵寄問卷為主，網路填報為輔，郵寄問卷有 3 大類(詳表 2.3.1)：

1. **行業別調查表**：大致以小類行業為區隔，全查層之大型雇主企業(具公司組織型態、所有多場所企業、部分單場所營業規模增長快，且薪資總額高於右截略點以上企業)使用長表(long form or census form)；抽查層之中型雇主企業(公司組織型態、單場所、薪資總額低於右截略點以下)使用短表(short form or sample form)。因此，非製造業各小類行業各有長、短表；製造業則因將年度抽樣調查(ASM；類似我國

工廠校正調查) 樣本併為抽查層樣本一部分，致使製造業各小類行業除長、短表外，另有 ASM 抽樣表，可視為短表之一部分，只是原先抽樣目的不同。

2. 公司組織結構表：供多場所企業填報。
3. 行業分類確認表：供規模小不需填行業別調查表，且經營行業別不明之企業填報。

表 2.3.1 美國經濟普查行業別問卷類型及合計數

行業類群	公司組織結構表	行業分類確認表	行業別調查表	合計
礦業		*	17	17
製造業		1	291	292
水電燃氣業、運輸、倉儲業		1	9	10
營造業		*	13	13
批發業		*	40	40
零售、住宿、餐飲業		7	33	40
金融、保險、不動產、租賃業		4	22	26
服務事業		18	65	83
多行業共用	1*	1*	0	2
島嶼區域	0	0	22	22
合計	1	32	512	545

註：「*」指共用調查表式。

不同於美國經濟普查，日本早年雖辦理企業及場所普查，惟其功能係以工商母體檔建置為目的，尚欠缺國家經濟帳所需之詳細財務問項，直至 2009 年方辦理類似美國之經濟普查；至南韓目前僅按分別辦理工業普查及服務業普查，並將於 2011 年辦理完整之經濟普查。

(二) 工商母體資料庫與各業別定期調查之整合體系：計 75 國答填此項，包括已開發國家如法國、德國、英國、澳洲、加拿大等 28 國，及開發中國家如俄國、巴西等 47 國，一般而言，採此法者仍以中、西歐及大英國協國家歷史較長，其中對整合經濟調查(Integrated Economic Surveys)體系統計過程有較多論述者，則屬加拿大及澳洲統計局。其作法係透過工商母體資料庫及各行業別定期經濟調查之整合體系，定期取得經濟統計資料。實言之，此法仍係經濟普查變革後之結果，因考量普查成

本及受查企業負擔等因素，遂採此法定期取得經濟統計資料。至經濟調查整合體系之調查方式大致如下：

1. 以工商母體資料庫為抽樣母體，由國家統計局統一抽樣。
2. 協調各辦理單位，調整自行辦理各業別調查之辦理時間、問卷內容、定義、樣本數，並排除不一致因素。
3. 未辦理抽樣調查之業別，協調業管單位，或國家統計局自行增辦抽樣調查，以期涵蓋全國經濟。
4. 資料處理及分析配合工商母體資料庫集中執行。

簡言之，此法係應用單一調查名冊、共同抽樣模式及問卷設計之理念、於每年(或每季、每月)定期辦理行業別或全經濟(economy-wide)之抽樣調查、且於調查後集中進行資料倉儲、處理(包括檢誤、插補、設算)、製表及分析作業。因於 5 年滾動週期(rolling five-year cycle)內取得完整資料，可以減輕集中一次辦理經濟普查之巨大工作負荷，且因與公務資料為主體之工商母體資料庫結合應用，可產生豐富之估計值，直接提供國家經濟帳(national accounts)及國內國民所得統計(GDP statistics)作業參用。

二、工商母體資料庫

工商母體資料庫向為美國辦理經濟普查或歐洲諸國辦理各行業別定期經濟調查時，掌握受查單位之重要依據。近 20 年來，各國國家統計局以此為主題，輪流主辦統計專題年會，其受重視程度可見一斑。因歐美國家彼此觀摩多年，各國工商母體資料庫之軟硬體雖不盡相同，但運作架構大致相近，發展策略方向亦趨一致，故以美國及義大利為例予以說明。

- (一) **美國**：美國工商母體資料庫係由企業場所名冊檔演進而來，其資料來源以稅務檔等公務資料檔為主，主要包括母體資料庫企業代號、企業名稱、地址、稅籍號碼 BAN、財務金額總數等資料項目，因其資料定期(每年或每季)更新，故亦為國家經濟

帳統計作業之資料基礎。其資料格式如表 2.3.2 所示。

表 2.3.2 美國普查局工商母體資料庫格式內容

欄位	項目內容
場所/企業代號系統	1. 普查場所代碼(CFN)：顯示現在所有權之 10 碼場所代號；前 6 碼是母企業代號
	2. 永久場所代碼(PPN)：不因所有權更替而變動之 10 碼場所代號。
	3. 依法設立之企業代號(EIN)：國稅局依法指定 9 碼代號為 EIN(相當我國企業之營所稅代號 BAN)，藉此向上連結母企業，向下連結所屬各場所。
企業組織名稱	包括法定登記成立之主要名稱及商業目的之次要名稱。
聯絡訊息	郵寄地址、實際地址、收件單位、主要填報人。
行業分類	參照 NAICS 標準分類，個別場所就主要業務，(單一或集團)企業依僱用員工最多之業務部分，分別進行行業分類。
營運分類	分辨批發業營運類別及管理場所類別，以補充「行業分類」之不足。
地區劃分	州、郡、市鎮、普查區、街廓。
組織法定類別	獨資、合夥、公司、合作協會、公有單位、其它(如基金會)。
企業規模及營運數據	1. 稅負狀況：可稅(營利單位)、免稅(非營利單位)。
	2. 員工人數：最近 3 年(含當年)資料(不包括派遣人力)，如有缺漏以迴歸模型設算，所以必須註明資料是實際的或經設算的。
	3. 第 1 季薪資(單位：千元)：包括場所及雇主企業當年第 1 季及前 2 年資料，如有缺漏以迴歸模型設算，所以必須註明資料是實際的或經設算的。
	4. 全年薪資(單位：千元)：由場所及法立企業各季資料合計之，如有缺漏以迴歸模型設算，所以必須註明資料是實際的或經設算的
	5. 各項收入(單位：千元)：包括場所及雇主企業之年度營業收入、利息收入、租金收入、毛收入及總收入；如有缺漏不進行設算。
	6. 資產(單位：千元)：包括場所及雇主企業之年度資產值；如有缺漏不進行設算。
其它資訊	營業、歇業、倒閉、售出等現況註記；所有權前後手等資訊；不同統計作業註記「該筆是否涵蓋在目標母體內」。

為維持母體資料庫品質，連結不同資料來源後，於比對及檢核過程如有必要，母體資料庫運作單位可就特定問題發出書面問卷或進行面訪，或參考年度經濟調查結果，惟多針對影響層面較深之中、大型企業集團進行；另亦針對多場所企業及企業集團之多層次控股權屬結構，參考經濟普查「公司組織結構表」之回表結果，除有助於母體資料庫之呈現，更可據以發布

企業生態結構統計(Enterprise Demography Statistics)，掌握企業生態之分布重點，對經濟政策有關鍵性之引導功能。

(二) 義大利：為持續發展重要產業，義大利國家統計局長期維護「行業別衛星資料庫」，其內容主要包括母體資料庫企業代號(主資料庫連結鍵值)、行業代號、稅籍號碼 BAN、特性別細項資料(業管公務資料為主，部分為經濟調查之檢核後結果)，其豐富之特性別細項資料除作為產業發展之前瞻工具，亦提供各項經濟調查重要之抽樣來源，如圖 2.3.1 及圖 2.3.2 所示。

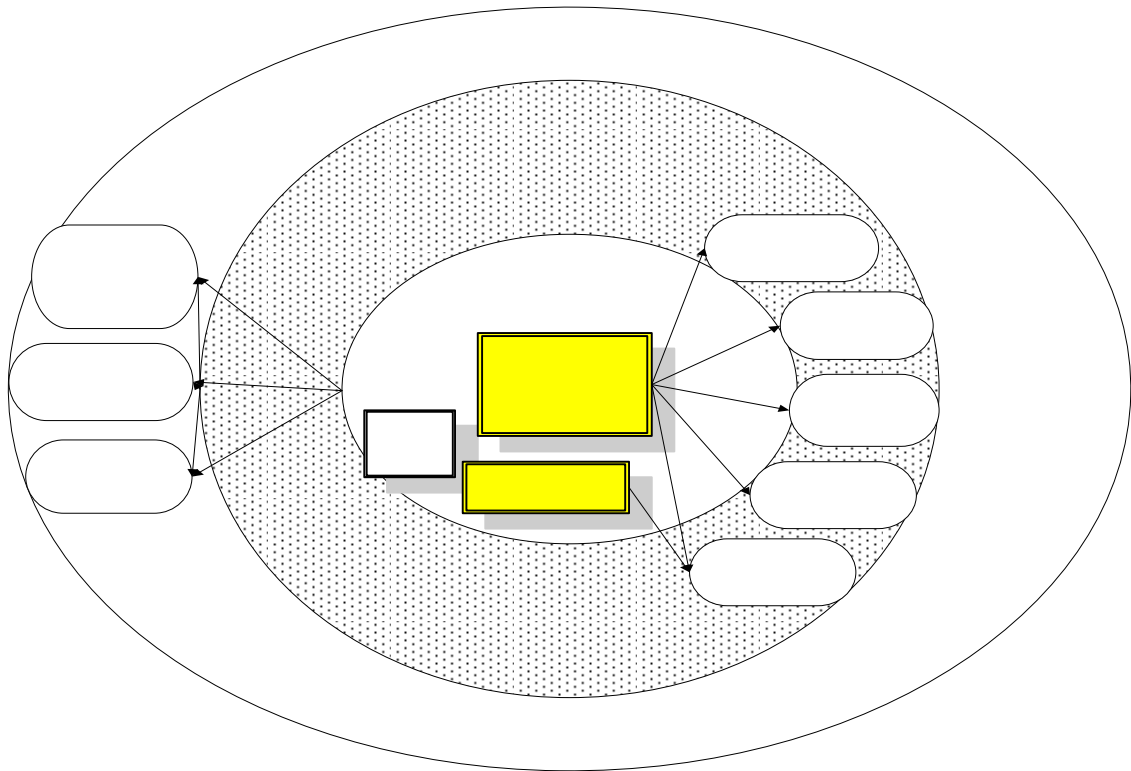


圖 2.3.1 工商母體資料庫－義大利

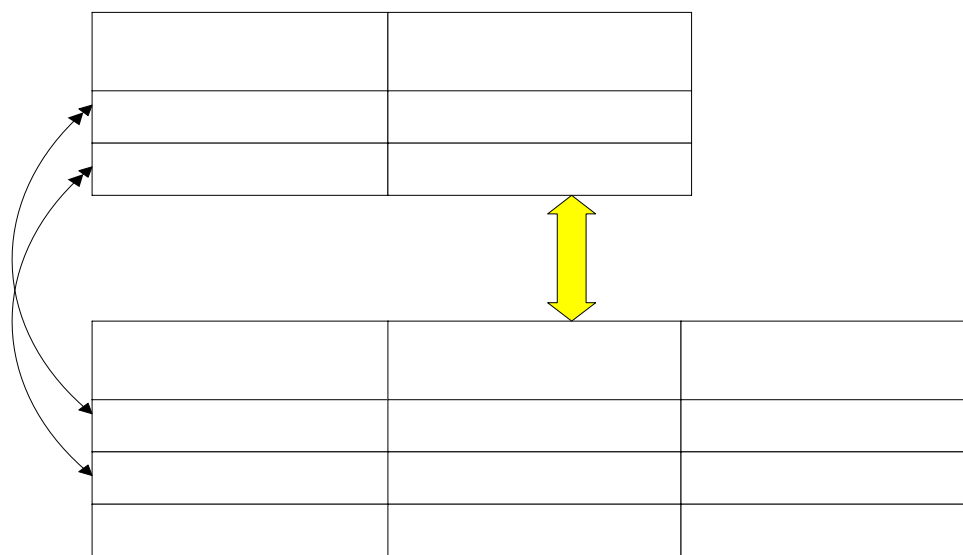


圖 2.3.2 工商母體資料庫及衛星資料庫關聯圖

企美

連結
相同
BAN

企美

第四節 各國經濟調查之抽樣及分層方式

一、**截略分層抽樣法**：各國辦理經濟普查或經濟調查時，在各項重要統計結果時間序列必須銜接，故各項調查必須賡續辦理之前提下，常囿於經費問題，而亟思降低抽樣規模之良方，其要務當屬提高抽樣設計成本效能。而依據「決定樣本大小(Determining Sample Size: Balancing Power, Precision, and Practicality, Patrick Dattalo, 2008)」一書之論述，一個有效能(effective)之抽樣設計，必須符合以下 4 項條件：

- (一)可達成研究目標。
- (二)可算出抽樣變異之準確估計值。
- (三)具備可行性。
- (四)達成研究目標之成本較低。

而前揭條件間可能存在衝突，故須從中取其平衡。

各國統計局一般應用於企業場所面抽樣調查，最具效能之抽樣設計當屬截略抽樣(cut-off sampling)。OECD (2006)對於截略抽樣之定義為：「抽樣母體中超過預定門檻值之場所單位納入調查樣本，低於此門檻值之場所單位則摒除之。」，其中之「門檻」值，通常與員工僱用或產出水準等衡量企業規模之變數有關。另 IMF(2004)曾評斷：「若截略抽樣之樣本摒除部分，造成之偏誤小，則它也是有效率之抽樣方法；如果要求的不是最小偏誤，而是最小均方差(mean square error)，即變異數(variance)與平方偏誤(squared bias)之和，截略抽樣亦不失為一個好選擇，因其變異數減低數可抵銷此偏誤，而尚有餘裕。」

另 James R. Knaub. Jr.(2008)在「截略抽樣及推論(Cutoff Sampling and Inference)」一文指出，截略抽樣產生之資料配適比率估計模型，即使迴歸線旁之資料點並非呈現均勻變異(heterodcedasticity or nonconstant variance around regression

line)，其估計值仍相當穩健，對未回表或不查層資料之插補亦有不錯之效果。故對許多場所面調查而言，當抽樣母體極度偏斜，但研究變數間迴歸關係之解釋變數穩定可得，截略抽樣結合比率估計不僅深具成本效能，且比其他可能替代方案更能減輕企業受查負擔，簡化調查過程，降低人力及成本，並產生準確之估計值。故美國普查局經濟普查、加拿大統計局及澳洲統計局之整合經濟調查均採此法辦理各項場所面統計調查。

自 1986 年以來，統計學家針對截略抽樣發展了諸多最適分層法(optimal size stratification)，Horgan, Kozak & Verma(2006)評估其中較佳者為 Hidiroglou(1986)及 Lavallee-Hidiroglou(1988)所提出的方法。兩者之明顯差異在於前者之副母體(subpopulation)僅分為「全查層」及「抽查層」2 類；後者則分為「全查層」、「抽查層」及「不查層」3 類。新加坡統計局(2008)之服務業經濟調查抽樣設計，特針對二法加以評估，由結果可知，如欲維持期望變異係數值(expected cv values)於 10%水準，Lavallee-Hidiroglou 法所需樣本抽出率可減低 19%，優於 Hidiroglou 法。

二、美國經濟普查之抽樣設計：美國經濟普查兼採全面調查及抽樣調查方式，其抽樣設計係以細類行業及地區為副母體，每副母體再依場所薪資支出為分層變數，以 2 個截略點區分全查層、抽查層及不查層，如表 2.4 所示。各層內涵如次：

表 2.4 截略點選取準則

行業類群	場所薪資 年度支出 右截略點	場所薪資 年度支出 左截略點	說明
製造、礦業	—	—	除不查層之外，製造業年度調查(ASM)樣本場所 58,400 家整合入普查以 ASM 表受查，非 ASM 大企業抽出 92,400 家以長表受查，非 ASM 中企業抽出 47,000 家以短表受查。各細業截略點不同。
水電燃氣、運輸、倉儲業	\$ 1,000,000	\$ 90,000	受查場所全查層約 138,000 家；抽查層抽出 6,700 家，抽出率約 9.5%。
營造業	—	—	受查場所規模以 10 人以下居多，全查層為多場所雇主約 12,000 家；抽查層為單

行業類群	場所薪資 年度支出 右截略點	場所薪資 年度支出 左截略點	說明
			場所雇主分 2 子層，比例配置，共抽出 110,000 家，抽出率約 20%。
批發業	—	—	全部納入全查層，受查場所約 407,000 家。
零售、住宿、餐飲業	\$ 1,000,000	\$ 28,000	受查場所全查層約 855,451 家，抽查層抽出 156,000 家，抽出率約 10%。
金融、保險、 不動產、租賃業	\$ 550,000	\$ 200,000	受查場所全查層約 335,000 家；抽查層抽出 23,000 家，抽出率約 9%。
服務事業	\$ 1,000,000	\$ 15,000	受查場所全查層約 1,050,000 家；抽查層抽出 156,000 家，抽出率約 9.5%。
註 1. 右截略點界分全查層及抽查層，左端截略點界分抽查層及不查層。			
註 2. 截略點係就各細類行業 (industry) 分別計算決定，本表呈現者係大類行業之綜合截略點。			

(一) **全查層**：包括具公司組織型態之多場所企業，以及營業規模增長快，且薪資總額高於右截略點之部分單一場所企業，不予抽樣，全數予以調查。

(二) **抽查層**：包括具公司組織型態之單一場所企業，且薪資總額低於右截略點者，本層通常再分為數個子層，進行簡單隨機抽樣，工業部門抽出率接近 20%，非工業則約為 9% 至 10%。

(三) **不查層**：單一場所且薪資總額低於左截略點之企業納入本層，全數不予調查，其資料以稅務檔等公務資料取代，惟經營行業別不明者須填報行業分類確認表。此類企業之規模特性一般為：從業員工 10 人以下，薪資合計占普查整體總薪資約 3%，收入合計占普查整體總收入約 10%，惟因家數眾多，為降低作業成本及填報負擔，爰採不予調查之方法。

三、 **經濟調查整合系統之抽樣設計及估計方法**：茲以加拿大及澳洲為例，介紹其經濟調查整合系統之抽樣設計及估計方法如次：

(一) 經濟調查整合系統之抽樣設計

1. **分層**：先按調查範圍，由工商母體資料庫擷取調查母體，再以小行業及地區進行第一階段分層，稱為自然層 (natural strata)；每一自然層內再進行第二階段分層，各子層稱為最終

層 (final strata)，包括全查層 (take-all stratum)、抽查層 (take-some stratum) 及不查層 (take-none stratum)，全查層及抽查層之臨界點，係以 Lavallee-Hidiroglou 法產生。

2. **區分不查層**：每一自然層內個別企業收入由大到小排序後，略去截略點 (cutoff point) 以下之小企業，即企業收入由小至大累計，低於自然層內總收入 10% 以下之小企業，歸入不查層。
3. **樣本配置**：依樣本變異及經費限制，進行樣本最適化配置，全查層可因專家建議另加入少數重要企業 (must-take)；抽查部分則依企業規模 (enterprise size) 變數 (如企業收入或員工人數) 分為 2 層。
4. **樣本抽取**：抽查層之各層以簡單隨機不置回法 (simple random selection, without replacement) 進行抽樣；並就工商母體資料庫使用之永久樣本序號 (permanent number)，以樣本協調 (sample coordination) 機制為原則，避免於不同年度發生重疊抽樣 (sample overlap) 現象，亦即不使相同廠商於鄰近調查週期納入樣本，俾降低填報負擔。

(二) 經濟調查整合系統之財務估計過程

1. **母體資料內涵**：工商母體資料庫資料範圍最為完整，惟資料項目較少且固定，僅含財務報表總額項目；整合經濟調查資料項目較多且較具變化，包括收支細項及經營特徵等，各項經濟調查因係由統計局統一抽樣，各調查年度樣本及工商母體資料庫之資料結構近似，爰可依財務資料總額與細額之比例關係，產生豐富之估計值。對個別企業資料，係採比率估計法或迴歸估計法推估；對於行業或地區別言，則採比例分配法 (pro-rating method) 予以彙總統計。另將工商母體資料庫之母體 (frame) 資料內容，依資料來源區分為僅直接調查、雙源收集、僅有稅務資料 3 類如次：

- (1) **僅直接調查 (Direct Collection only)**：簡稱 **D** 群企業資料，因含複雜企業結構或事業羣之受查單位，無法與企業統一

編號一對一連結，本類企業單位數比重約占 18%，營業所得比重則占 63%。

(2) **雙源收集** (Both Direct and Tax Collection)：簡稱 **B** 群企業資料，包括單純企業結構之被調查單位，可藉企業統一編號一對一連結，本群企業單位係經濟活動調查之抽查對象，調查資料可與營業所得稅結合進行推估，其企業單位數比重約占 24%，營業所得比重則占 26%。

(3) **僅有稅務資料** (Tax Data only)：簡稱 **T** 群企業資料，包括多數小型企業，可藉企業統一編號一對一連結，本群企業單位數比重約占 58%，營業所得比重則占 11%。因稅務資料僅為財務資料總額，缺少細項財務金額，故須借用 **B** 群企業資料之細項比例結構，推估本群企業資料之細項財務金額。

2. **推估**：營業所得稅檔中 **B** 群、**T** 群企業資料缺漏部分，多屬停歇業企業，故經清除停歇業者，並迴歸插補少部分尚存續者後，營業所得稅檔趨於完整，調查資料方可與營業所得稅結合進行後述推估。其中 **B** 群企業資料係採比例分配法 (pro-rating method) 結合調查資料可與營業所得稅進行推估：

假設 $t_{\text{調查,企業}i}$ 是 i 企業經濟活動調查之財報資料總額， $c_{\text{調查,企業}i}$ 是 i 企業經濟活動調查之財報資料細項； $t_{\text{稅務,企業}i}$ 是 i 企業營業所得稅檔之財報資料總額；小寫 t, c 是個別企業之數值，中(小)業之財報資料總額及細項分別以 \hat{T} 及 \hat{C} 表示，則 **B** 群企業之 h 中行業、 k 小行業財報資料細項之比例分配估計值為：

$$\hat{C}_{小業k}^{B群} = \frac{\hat{C}_{調查,中業h}^{B群}}{\hat{T}_{調查,中業h}^{B群}} * T_{稅務,小業k}^{B群} \text{-----} (a)$$

為確保財務細額估計值各 $\hat{C}_{小業k}^{B群}$ 可加總至財報資料總額，採 $T_{稅務,小業k}^{B群}$ 為財報資料總額控制值(control values)。如將公式延伸應用至 T 群企業資料，財務細額之比例分配估計式為：

$$\hat{C}_{小業k}^{T群} = \frac{\hat{C}_{調查,中業h}^{B群}}{\hat{T}_{調查,中業h}^{B群}} * T_{稅務,小業k}^{T群} \text{-----} (b)$$

即 T 群之比例分配因子(pro-rating factors)係借 B 群之比例

$\frac{\hat{C}_{調查,中業h}^{B群}}{\hat{T}_{調查,中業h}^{B群}}$ ，雖可能增加 T 群企業資料估計之變異及偏誤，惟因

T 群企業收入占全體企業收入低於 15%，影響層面不深，故可接受。綜合上述，經濟調查的估計方法實為內插、估計、外插三步驟之合成：

- (1) 內插：先以迴歸或其它模型，插補營業所得稅檔缺漏之個別企業財報資料總額。
- (2) 估計：在調查樣本($c_{調查,企業i}, t_{調查,企業i}$)點為基礎之二維迴歸資料散布圖內，針對抽查層內非調查樣本之企業，以下列估計式(c)，進行比率估計。

$$\hat{C}_{企業i} = \hat{r}_{某自然層之抽查層} * t_{稅務,企業i} + \hat{\varepsilon}_{企業i} \text{-----}(c)$$

其中 $\hat{r}_{\text{某自然層之抽查層}}$ 是某自然層之抽查層中各 $(c_{\text{調查,企業}i}, t_{\text{調查,企業}i})$ 點所構築迴歸直線之斜率，且截距為零。再依行業別加總企業比率估計值，即可估計行業別財報資料細項合計，並取代前揭(a)，(b)式之直接比例分配估計之行業別財報資料細項，且二種方式估算之行業別財報資料細項估計值可相互核對，至營業所得稅檔財報資料總額，則可為其下細項估計值之控制值。

(3) 外插：借用 **B** 群(抽查層)企業資料之財務細項分配結構，外插(extrapolation) **T** 群(不查層)企業資料之財務細項。

整體而言，整合經濟調查之估計過程係結合稅務檔缺漏內插，抽查層迴歸(比率)估計，及不查層借相同自然層之抽查層之比率估計模式，進行外插的一種合成估計(synthetic estimation)。至各自然層中之不查層，其估計比率則藉臨近抽查層比率推計之。

第五節 各種分層方式比較

因各行業經營型態及財務結構相異，將普查資料加以分類，使一群內資料之同質性愈高，不同群之相異度愈高，於群內樣本統計量推估母數時，可更為精確，故分層效果愈佳，愈能精確反映母體。

傳統集群分析(cluster analysis)方法可分成階層式(hierarchical method)及非階層式(nonhierarchical method)兩大類，而階層式又分成集結法(agglomerative method)及分裂法(divisive method)兩種。除傳統集群分析外，常見分類方法尚包括模糊集群分析(FCM Fuzzy C-Means)、自動互動檢視法(CHAID Chi-square Automatic Interaction Detection)及品質門檻(Quality Threshold)分群法等(詳表 2.5)。茲列述介紹如次：

一、 階層式集群分析：

(一) 集結法：本法係先計算各觀察值的距離或組內誤差矩陣，將最接近的兩個觀察值合併成同一集群，接著再算出合併後的觀察值距離或組內誤差矩陣，並重複以上步驟，直到所有觀察值合併成同一集群。其方法較常使用者有單一連結法(single linkage method)、完全連結法(complete linkage method)、平均連結法(average linkage method)、形心法(centroid method)及華德法(Ward's method)。

(二) 分裂法：本法概念恰與前法相反，係依序將相異性最大之觀察值分裂成另一集群，步驟如下

1. 首先將所有觀察值視為一個集群。
2. 計算哪一個觀察值與集群的相異性最大或相似性最小，然後將其分裂成另一集群。
3. 分別計算大集群中每個觀察值與集群內及集群外的平均距離，若比較接近另一個集群，則將此觀察值分配到另一集群，否則就留在集群內。重複此過程直至兩個集群內的觀察值都無法再分裂為止。

4. 重複第 2、3 步驟，將所有觀察值一直分裂到單一集群為止。

二、 **非階層式集群分析**：k 平均法，步驟如下(Sharma 1996)：

(一) 預先假設 k 個集群數，k 個種子點。

(二) 計算各觀察值到各集群形心的距離，將其分配至最近一群。

(三) 重新分配每個觀察值到適當的集群，直到達到事前訂定的停止標準。

(四) 重複(二)、(三)步驟，直到無法重新分配為止。

三、 **模糊集群分析法**：加州大學教授 Zadeh(1965)首先提出模糊集合理論，並將隸屬度(Membership)及隸屬函數(Membership Function)的概念引入了傳統集群分析法中，其中隸屬度表示某物體屬於某集合「模糊程度」的大小。至本法最早係由 Dunn(1974)以傳統的 k 平均法，結合模糊理論而成，稱之為「模糊化 k 平均法」。另 Zimmermann(1991)認為事物的模糊性可分為兩種，一種為本質上的模糊，一種為資訊上的模糊，前者常隨情境的不同而有差異，後者則因資訊的不足或內容模糊而形成模糊性。Leung Yee(1984,1985)亦曾以模糊理論探討臨界值分類方法。Rousseeuw (1987)提出以側影圖(Silhouette)判定分群之效果，並提出側影係數 SC(Silhouette Coefficient)輔助，以求得最佳分群數，並以目標函數來判斷元素與中心點的相近度，而且在尋找最佳相近度時，採用最小目標函數準則；目標函數事實上就是模糊集群分析轉化最小平方法(least squared error method)的誤差平方和函數(square error function)。

四、 **品質門檻分群法**：由 Heyer, Kruglyak, Yooseph 於 1999 年所提出，此法需設定 2 個門檻值，第 1 個為歸屬同群的資料相關性標準，第 2 個為分群品質標準，即群內最少資料筆數的門檻值。分群結果必須同時符合 2 個門檻值，惟仍可能有觀察值無法被有效分群。

五、**自動互動檢視法**：本法最早由 Sonquist 和 Morgan 於 1960 年代初期提出，稱為 AID(automatic interaction detection)，係利用逐次搜尋的方式，繪出樹狀圖來找出最能解釋準則變數變化的預測變數；分析資料則採無母數方法，故並未預設資料之分布。一般而言，本法係採 Bonferroni 調整卡方值作為分割樣本的依據，利用卡方檢定，使同質的樣本單位能歸於同一群體，接續使用半層次(semihierarchical)逐次搜索方法，對每一預測變數的反應水準進行成對、合併及分割，以求得各預測變數水準的最少分群數。一旦分群數目確定，再以顯著性最大的預測變數將原始樣本分割成若干小群體；而每一小群體又依同樣方式繼續分割，直至分群結果無顯著差異，或分群後樣本數過小而停止。在 Answer Tree 套裝軟體中，合併變數方式會因資料類型而有所差異。若屬順序尺度之變數，僅以其資料順序合併；反之若屬類別尺度之變數則無上述限制。另解讀資料分析方面，該軟體可以條件機率解釋所有具從屬關係之父節點(parent nodes)與子節點(child nodes)的關聯性。

表 2.5 集群分析

分層方法		分群方式	優點	缺點
階層式	集結法	單一連結法	1. 不需事先確定分成幾群或種子點。 2. 傳統集群分析明白容易了解。 3. 資料屬混合多變量較精確。	1. 觀察值分群後即不再變動。 2. 距離的定義不同，分群的結果可能不同。 3. 易受連結效應影響，觀察值傾向指派給已存在的分群，而不是另外新增群。 4. 未考慮變數間互動關係。
		完全連結法		
		平均連結法		
		形心法		
		華德法		
	分裂法	係依序將相異性最大之觀察值分裂成另一集群。		
非階層式	k 平均法	分配每個觀察值到適當的集群，直到達到事前訂定的停止標準。	觀察值分群後，仍會變動。	1. 需事先確定分成幾群、種子點。 2. 未考慮變數間互動關係。

分層方法	分群方式	優點	缺點
模糊集群分析	以目標函數來判斷元素與中心點的相近度，而且在尋找最佳相近度時，採用最小目標函數的準則。	<ol style="list-style-type: none"> 1.分群後，還能保有原始資料的訊息，可供後續之研究。 2.觀察值可明確分群。 3.以客觀指標判斷分群效果。 	<ol style="list-style-type: none"> 1.因為模糊集群分析法是考量整體狀況加以分群，所以每個集群中，缺乏代表性的樣本。 2.計算費時。 3.未考慮變數間互動關係。
QT 分群法		<ol style="list-style-type: none"> 1.適用於醫學上嚴謹基因 (gene) 挑選方法。 2.以觀察值的相關係數，而非距離測量同一群的變異程度。 3.不需事先確定分成幾群。 4.即使重複執行，分群結果仍相同。 	<ol style="list-style-type: none"> 1.仍有可能觀察值未分群。 2.計算費時。 3.未考慮變數間互動關係。
自動互動檢視法	半層次的逐次搜索方法，對每一預測變數的反應水準進行成對的合併然後分割，以求得各預測變數水準的最少分群數。	<ol style="list-style-type: none"> 1.考慮變數間互動關係。 2.以逐次搜尋的方式，繪出樹狀圖，易於了解。 3.可即時初步過濾重要變數。 4.可適用類別資料及連續資料。 5.事先確認分層的觀察值數目。 6.快速產生結果，易了解層內結構與各變數關係。 	<ol style="list-style-type: none"> 1.樣本數 2000 個以上才穩定。 2.未考慮變數相關問題，分析過程中，變數出現的次序並不表示其重要性。 3.偏態分布的變數可能不出現於樹狀圖。 4.訂定之停止規則使分割過早停止，可能忽略重要變數。

第三章 研究方法

工商普查資料因各行業營運狀況不同，抽樣調查部分爰設計製造業、營造業、商業、運輸及倉儲業、金融及保險業、服務事業及補習教育及托育服務業等計 7 種調查表式，復依財務問項多寡，區分抽樣調查甲表及乙表，並依據前節之抽樣方法，將抽出之廠商依業別查填正確表式。茲以 95 年普查資料為例，運用以下統計方法分析其抽樣效能：

一、 **研究對象與範圍**：95 年普查母體企業單位計 110 萬 5 千家，樣本計 7 萬 3 千家(其中縣市調查約 4 萬 8 千家，餘為分工單位調查對象)，將資料依行業特性區分 18 大類。

二、 **資料分析方法**：

(一) **敘述性統計**：觀察母體與樣本家數分布、分層情形，了解整體普查資料特性、樣本配置及分層效果，並分析樣本空層原因。

(二) **關聯分析**：普查資料包括類別資料及連續資料，本部分係藉關聯係數呈現變數間之關聯程度，進而分析樣本結構。依變數性質大致分為三個部分：

1. **兩連續變數**：兩連續變數係以相關係數 $\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$ 衡

量其相關程度，可能為正相關或負相關，當兩變數呈正相關時，則一變數高數值傾向於對應另一變數之高數值，而一變數低數值傾向於對應另一變數之低數值；負相關則相反，一變數高數值傾向於對應另一變數之低數值，而一變數低數值傾向於對應另一變數之高數值。 ρ^2 意義為某一變數已知時，另一變數的變異減少百分比，也就是當用某一變數預測另一

變數時，可減少 $\rho^2 \times 100\%$ 的錯誤。

2. **兩類別變數**：兩類別變數係以「消減錯誤」之比例大小衡量，消減錯誤係指某特徵值以輔助變數觀察時，可降低之預測錯誤，相關程度愈高，則預測錯誤愈少，一般以 PRE (Proportionate Reduction in Error) 測量法計算。假設衡量兩變數 X 、 Y 之相關程度，其意義係指運用 X 預測 Y 能夠減除錯誤的比例。

(1) **Goodman and Kruskal Lambda 關聯測量法**：以眾數作為預測準則，資料屬對稱分配時，兩變數可互相預測，以衡量消減錯誤之比例。其內涵如次：

X 未知，預測 Y 時的錯誤機率：

$$P(\varepsilon_1) = 1 - \frac{1}{2} \frac{M_y}{N} - \frac{1}{2} \frac{M_x}{N} = 1 - \frac{M_y + M_x}{2N}$$

X 已知，預測 Y 時的錯誤機率：

$$P(\varepsilon_2) = 1 - \frac{1}{2} \sum \frac{m_y}{N} - \frac{1}{2} \sum \frac{m_x}{N} = 1 - \frac{\sum m_y + \sum m_x}{2N}$$

即當 X 分類之各值對 Y 預測所得之資訊，成功預測機率為

$$\frac{1}{2} \sum \frac{m_y}{N},$$

$$\lambda = \frac{P(\varepsilon_1) - P(\varepsilon_2)}{P(\varepsilon_1)} = \frac{\text{消減之誤差量}}{\text{不正確預測之機率}}$$

$$\lambda = \frac{\left(1 - \frac{M_x + M_y}{2N}\right) - \left(1 - \frac{\sum m_x + \sum m_y}{2N}\right)}{1 - \frac{M_x + M_y}{2N}} = \frac{\sum m_x + \sum m_y - (M_x + M_y)}{2N - (M_x + M_y)}, \text{ 其中}$$

$M_y = Y$ 變項的眾數值 (眾數所在對應之發生次數)。

$M_x = X$ 變項的眾數值。

$m_y = X$ 變項的每個值之下， Y 變項的眾數值。

$m_x = Y$ 變項的每個值之下， X 變項的眾數值。
 $N =$ 總數。

(2) **Goodman and Kruskal Tau 關聯測量法 (不對稱)**: 係將邊際次數取代眾數，避免分布過於集中，導致影響測度的偏差。

$$X \text{ 未知，預測 } Y \text{ 時的總錯誤： } E_1 = \sum_{j=1}^r \frac{(N - F_{y_j}) \times F_{y_j}}{N}$$

X 已知，預測 Y 時的錯誤：

$$E_2 = \sum_{j=1}^r \sum_{i=1}^c \frac{(F_{x_i} - F_{y_j x_i}) \times F_{y_j x_i}}{F_{x_j}}, \text{ 其中}$$

$$\frac{(F_{x_i} - F_{y_j x_i}) \times F_{y_j x_i}}{F_{x_j}} : X \text{ 已知，預測 } Y \text{ 時未解釋的次數}$$

$F_{y_j x_i}$: 權數

$$\text{tau} = \frac{E_1 - E_2}{E_1}, \text{ 其中}$$

$Y_j = Y$ 變項值

$X_i = X$ 變項值

$F_{y_j} = Y$ 變項之邊際次數， $F_{x_i} = X$ 變項之邊際次數

$F_{y_j x_i} =$ 同屬於 Y 的值 j 和 X 的值 i 之個案數目

$r = Y$ 變項值的數目， $c = X$ 變項值的數目， $N =$ 總數

3. **類別變數與連續變數**: 如兩變數分別為類別變數與連續變數，則以 ETA(E) 係數值衡量，係以一個類別變數 (X) 為自變項，一個連續變數 (Y) 為依變項，根據自變項的每一個值，

預測或估計依變項之平均數。ETA 的平方值(E)² 表示消減誤差的比率。

$$E^2 = \frac{\sum_{i=1}^m \sum_{j=1}^N (Y_{ij} - \bar{Y})^2 - \sum_{i=1}^m \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2}{\sum_{i=1}^m \sum_{j=1}^N (Y_{ij} - \bar{Y})^2} = E^2 = \frac{\sum N_i \bar{Y}_i^2 - N \bar{Y}^2}{\sum Y^2 - N \bar{Y}^2}, \quad E = \sqrt{E^2}, \quad \text{其中}$$

Y_{ij} ：依變項的數值

\bar{Y} ：依變項的平均數(overall mean)

\bar{Y}_i ：每個自變項值(X_i)上的平均數。

N 為全部個案數目。

N_i 為每個自變項值(X_i)的數目。

由於依變項(Y)屬連續變數，若自變項(X)未知，以平均數作為預測的準則，全部誤差為 $\sum (Y - \bar{Y})^2$ 。若自變項已知，其預測值應為對應自變項值 X_i 的各依變項 Y_{ij} 的平均數(\bar{Y}_i)；預測之誤差總數為 $\sum_{i=1}^m \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2$ ，所消減的誤差為 $\sum_{i=1}^m \sum_{j=1}^N (Y_{ij} - \bar{Y})^2 - \sum_{i=1}^m \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2$ ，藉由 X 之解釋，可消減錯誤之比率。由此可知 ETA 係以均值作預測的準則。

(三) **樣本精確度分析**：一般而言，企業規模之分布，皆有偏斜之現象，根據 95 年普查資料顯示，營業收入未滿 500 萬元之企業家數占 67.63%，其生產總額卻僅占 4.40%；至營業收入 5 億元以上之企業僅占 0.61%，其生產總額卻達 66.68%，因資料嚴重偏斜，無法僅以平均數衡量樣本之精確度，爰運用截尾平均數(trimmed mean)法(黃登源，2003)，針對人數、營業收入及資產淨額，先將偏斜資料轉換後，計算截尾比例，再導出平均數、變異數、偏態係數及峯態係數等 4 項動差，計算原始資料的平均值，比較各細業樣本截尾平均數與母體截尾平均數，期

以此衡量樣本之精確度。

(四) **CHAID 分析**：普查資料偏斜之特質，非常適合運用 CHAID 予以分析，該分析係應用無母數方法，分析並解釋資料的結構及分布，分析前毋須假設資料分配，且可與其他統計方法互補，因而被廣泛應用。因整體工商業母體雖可掌握，卻無法將廠商營運資料予以有效分類，以了解結構性的轉變；復以廠商多角化經營日趨普及，以新興管道方式(如透過網際網路)獲取高利潤之廠商，致人力創造利潤之概念式微，如單以營業收入觀察產業變化，恐已不符合現況，爰將數個自變數同時納入考量，並觀察變數間之關係。而分析時若僅考慮自變數間的相關程度，可能忽略自變數間亦存在互動，且此互動可能影響自變數與營業收入(準則變數)之關係，故應用此法可同時觀察所有可能影響營業收入的因素，並找出其間互動。

(五) **分類及迴歸樹(Classification and Regression Tree, CART)分析**：本法與 CHAID 同屬決策樹分析法，最初被應用於醫學診斷及預測上，近年應用之範圍則日趨廣泛。CART 係一種透過分割技術，將觀察值以二元及序列展開方式進行分類，並降低各類別分散程度之無母數分析工具；且其對於根節點至子節點上之同一變數可重複檢驗，亦可透過持續分割，達到最佳分類組合。分析流程如下：

1. 分類樹之建立：

(1) 分割之選擇：CART 之資料分割原則，係透過每一個資料觀測值之分割，計算所產生子節點之不純度(impurity)統計量，從而找出減少最大不純度的分割方式，完成分割點之選取。

(2) 繼續分割或視為最終節點之準則：當節點中所有屬性變數之值皆相同，或節點中樣本數已達預先設定時，則停止分割，否則繼續分割。

(3) 將最終節點予以分類。

2. **初步變數選取方式**：分析時，將初步衡量自變數在節點上的重要性，進而篩選較重要之自變數進一步分析。考慮在節點 t 分割成子節點 t_L 及 t_R 時，存在下列關係：

$$\begin{aligned}
\sum_{x_n \in t} (y_n - \bar{y})^2 &= \sum_{x \in t_L} (y_n - \bar{y})^2 + \sum_{x \in t_R} (y_n - \bar{y})^2 \\
\sum_{x_n \in t_L} (y_n - \bar{y})^2 &= \sum_{x \in t_L} (y_n - \bar{y}_{t_L})^2 + \sum_{x \in t_L} (\bar{y}_{t_L} - \bar{y})^2 \\
\sum_{x_n \in t_R} (y_n - \bar{y})^2 &= \sum_{x \in t_R} (y_n - \bar{y}_{t_R})^2 + \sum_{x \in t_R} (\bar{y}_{t_R} - \bar{y})^2 \\
SST_t &= \sum_{x_n \in t} (y_n - \bar{y})^2 = \sum_{x \in t_L} (y_n - \bar{y})^2 + \sum_{x \in t_R} (y_n - \bar{y})^2 \\
&= \left\{ \sum_{x \in t_L} (y_n - \bar{y}_{t_L})^2 + \sum_{x \in t_R} (y_n - \bar{y}_{t_R})^2 \right\} + \left\{ N(t_L)(\bar{y}_{t_L} - \bar{y})^2 + N(t_R)(\bar{y}_{t_R} - \bar{y})^2 \right\} \\
&= SST_t + SSB_t
\end{aligned}$$

其中 Y 為準則變數， X_i 為自變數，整體資料筆數為 N ， $N(t)$ 為節點 t 的樣本數，以

$$\hat{C}_{X_i}(t) = \frac{\frac{1}{N(t)} SSB_{X_i}(t)}{\frac{1}{N} SST} = \frac{\frac{N(t_L)}{N(t)} (\bar{y}_{t_L} - \bar{y})^2 + \frac{N(t_R)}{N(t)} (\bar{y}_{t_R} - \bar{y})^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}, i = 1, \dots, k.$$

衡量節點 t 時， X_i 為分割變項時的重要性， $\hat{C}_{X_i}(t)$ 越大， X_i 在節點 t 的重要性越大；以整體分類樹來看，亦以 $\hat{C}_{X_i}(t)$ 來衡量路徑的重要。

3. **變數評估**：初步篩選自變數後，重新進行 CART 分析，並以 I_{X_i} 及 $R_{x_i}^2$ 兩統計量衡量移除 X_i 變項後，對分割的影響，計算方式如下：

(1) 移除 X_i 變項，節點 t 分割成子節點 t_L 及 t_R 時

$$I_{X_i}(t) = \frac{\frac{N(t_L)}{N(t)}(\bar{y}(t_L) - \bar{y})^2 + \frac{N(t_R)}{N(t)}(\bar{y}(t_R) - \bar{y})^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}, i = 1, \dots, p.$$

設 \tilde{T}_i 為表示所有的中間節點， $\#(\tilde{T}_i)$ 為所有中間節點的個

數，可得

$$I_{X_i} = \frac{1}{\#(\tilde{T}_i)} \sum_{t \in \tilde{T}_i} I_{X_i}(t)$$

I_{X_i} 值越小，表示在移除 X_i 變項後，所有中間節點對 SSB

的加權平均變異的影響越小，亦表示 X_i 變項越重要。

(2) 以 $R_{X_i}^2$ 表示在移除 X_i 變項後， SSB 對總變異 SST 的解釋比例：

$$R_{X_i}^2 = \frac{\sum_{t \in T_i} SSB(t)}{SST} = \frac{\sum_{t \in T_i} N_t (\bar{y}(t) - \bar{y})^2}{\sum_{j=1}^N (y_j - \bar{y})^2}$$

其中 T_i 表示在移除 X_i 變項後所有的最終節點，則 $R_{X_i}^2$ 值愈小，亦表示 X_i 變項愈重要。

(六) 對數線性模式 (Log-Linear Model) 分析：對數線性模式分析可了解各種不同變數的組合，其次數分布所反映之樣本特色與訊息。以普查樣本次數分布觀察，欲了解廠商經營模式轉變及產業變遷，必須先了解家數受何變數影響較大，方可剖析整體產業間的變遷，且本分析可觀察多個變數在某一水準下，發生次數的重要影響因素，對項目多元之普查資料而言，可發現重要變數間是否隱藏廠商分布之重要資訊，對於產業結構變化之了解可更為深入，從而提升抽樣資料分析之運用價值，並呈現 95 年普查抽樣方法之效能情形。

(七) 對應分析(Correspondences Analysis)：主要在於處理兩個或多個類別變項之間的關係，強調空間分布圖顯示資料結構的特徵：

1. 以列聯表為基礎將資料的特性呈現於空間分布的圖示中，需要幾何學及矩陣運算之數理理解。透過各個類別在空間的分布距離探討類別間的關係，將複雜的資料結構視覺化。
2. 不在於檢證資料是否適配研究者設定的理論模型，而係呈現資料本身的結構。座標軸標上分類點，以列或行分類之分數構建空間分布圖；經分數常態化(normalization)後，列或行點間歐氏空間距離(euclidian space distances)與卡方(chi-squared distances)距離相等。列或行點間之聚散情形可由其間卡方距離遠近來描述，當某 2 列(或行)點距離近，表示特徵相似(similar)；反之，其特徵差異大，則距離遠。

對應分析是列聯表資料的加權主成分分析(Principal Components Analysis)或因素分析(Factor Analysis)，用它去尋求列聯表的行列變數之間聯繫的低維圖示法。此方法的關鍵係利用一種資料變換方法，使含有 n 個樣本 m 個變項的原始資料矩陣變成另一個矩陣，並使 $R=Z'Z$ (分析變項之間關係的共變異矩陣 (Covariance matrix)) 與 $Q=ZZ'$ (分析樣本之間關係的共變異矩陣) 具有相同的非零特徵根，它們相應的特徵向量之間也有密切的關係。對共變異矩陣 R 、 Q 進行加權主成分分析或因素分析，分別能提取兩個最重要的公因數 R_1 、 R_2 與 Q_1 、 Q_2 。由於採取的是一種特殊變換方法，公因數 R_1 與 Q_1 在本質上是相同的，同理， R_2 與 Q_2 在本質上也是相同的，故可用 dim_1 作為 R_1 、 Q_1 的統一標誌；用 dim_2 作為 R_2 、 Q_2 的統一標誌，於是可將 (R_1, Q_1) 和 (R_2, Q_2) 兩組數據點，在由 (dim_1, dim_2) 組成的同一個直角坐標系中呈現。以便於觀察變項與樣本之間的相互關係。

奇異值(singular value)為特徵值(eigenvalues)之開方

值。特徵值為分析中變項分類間典型相關係數(canonical correlation)之最大值。慣性量(inertia)屬力學名詞，係指每一物件都有一重力中心(a center of gravity)或質心(centroid)，物件中每一粒子(particle)皆具質量 m 與質心 d 距離，因此其慣性量 $= md^2$ 。推演至對應分析，當一群集(cloud)剖面點(profile points)或變項分類(category)個別擁有可合計 1 之不同權值(masses)，為計算剖面點距離之加權係數：

第 i 列剖面權值 $= \frac{\text{列聯表第 } i \text{ 列邊際次數}}{\text{列聯表總次數}}$ ，如表 3.1：

表 3.1 第 i 列剖面權值

列	行				列剖面權值
	1	2	...	j	
1	$n_{11}/n_{.1}$	$n_{12}/n_{.2}$...	$n_{1j}/n_{.j}$	$n_{1.}/n_{..}$
2	$n_{21}/n_{.1}$	$n_{22}/n_{.2}$...	$n_{2j}/n_{.j}$	$n_{2.}/n_{..}$
...	
i	$n_{i1}/n_{.1}$	$n_{i2}/n_{.2}$...	$n_{ij}/n_{.j}$	$n_{i.}/n_{..}$
合計	1	1	1	1	1

第 j 行剖面權值 $= \frac{\text{列聯表第 } j \text{ 行邊際次數}}{\text{列聯表總次數}}$ ，如表 3.2

表 3.2 第 j 行剖面權值

列	行				列剖面權值
	1	2	...	j	
1	$n_{11}/n_{1.}$	$n_{12}/n_{1.}$...	$n_{1j}/n_{1.}$	1
2	$n_{21}/n_{2.}$	$n_{22}/n_{2.}$...	$n_{2j}/n_{2.}$	1
...	1
i	$n_{i1}/n_{i.}$	$n_{i2}/n_{i.}$...	$n_{ij}/n_{i.}$	1
合計	$n_{.1}/n_{..}$	$n_{.2}/n_{..}$...	$n_{.j}/n_{..}$	1

則每一列剖面點之慣性量可由距離之加權係數乘以與質心相距卡方距離(chi-square distance)，即 $m_i d_i^2$ 計算。因此，當 $r_{ij} = n_{ij}/n_{i.}$ ， $\bar{r}_j = n_{.j}/n$ ； $c_{ij} = n_{ij}/n_{.j}$ ， $\bar{c}_i = n_{i.}/n$ ； $p_{ij} = n_{ij}/n$ 時，

第 i 列剖面慣性量 (Inertia for the i th row profile) =

$$m_i d_i^2 = m_i \sum_j \frac{(r_{ij} - \bar{r}_j)^2}{\bar{r}_j}$$

其中列剖面權值為 m_i ，列剖面點與質心相距卡方距離

$$= d_i^2 = \sum_j \frac{(r_{ij} - \bar{r}_j)^2}{\bar{r}_j}$$

第 j 行剖面慣性量 (Inertia for the j th column profile) =

$$m_j d_j^2 = m_j \sum_i \frac{(c_{ij} - \bar{c}_i)^2}{\bar{c}_i}$$

其中行剖面權值為 m_j ，列剖面點與質心相距卡方距離

$$= d_j^2 = \sum_i \frac{(c_{ij} - \bar{c}_i)^2}{\bar{c}_i}$$

$$\text{總慣性量 (Total Inertia)} = \sum_i \sum_j \frac{(p_{ij} - r_i c_j)^2}{r_i c_j}$$

慣性量分配比 (Proportion of Inertia Accounted for) 係指個別變項解釋變異量 (variance) 占模型整體解釋變異量之分配比率，即前述總慣性量呈現之總變異，有多少出現在列剖面或行剖面之分配比率。

第四章 資料分析

第一節 母體資料分析

壹、95 年普查母體資料

95 年普查係依據「中華民國行業標準分類(第 8 次修訂)」區分 18 個大類、82 個中類、223 個小類及 495 個細類，除運輸及倉儲業中個人計程車係參考交通部相關公務及調查檔案資料推估外，餘母體家數計 1,065,250 家，並提供 7,510,060 個就業機會，投入 1,011,069 億元資產，創造 439,996 億元營業收入(詳表 4.1.1)。各重要統計結果分布情形如次。

- 一、**家數**：95 年普查母體家數逾 106 萬家，以批發及零售業 49 萬家最多，占 46.00%；製造業近 15 萬家次之，占 13.90%；住宿及餐飲業、其他服務業及營造業所占比率依序為 8.33%、7.67%及 6.89%；其餘各大業母體家數分配比皆小於 5.00%。
- 二、**從業員工**：95 年普查從業員工人數約 751 萬人，以製造業近 270 萬人最多，占 35.90%；批發及零售業僅近 189 萬人次之，占 25.17%；營造業與金融及保險業分別為 48 萬及 38 萬人，分別占 6.38%及 5.02%；其餘各大業從業員工人數分配比皆小於 5.00%。
- 三、**資產淨額**：95 年普查自有資產淨額為 101 兆元，金融及保險業因流動資產龐大，以 59 兆元居冠，所占比率高達 58.64%；製造業 21 兆元次之，占 20.95%；批發及零售業近 6 兆元，占 5.73%；其餘各大業自有資產淨額分配比皆小於 5.00%。
- 四、**營業收入**：95 年普查營業收入近 44 兆元，以製造業近 20 兆元最多，占 44.83%；批發及零售業 11 兆元次之，占 26.13%；金融及保險業 6 兆元，占 12.92%；其餘各大業營業收入分配比皆小於 5.00%。

製造業與批發及零售業除自有資產投入所占比率較低之外，其餘家數、從業員工人數及營業收入，兩大行業皆分居前兩名，其他各大

行業所占比率雖較低，但其分配比率之排名變動極大，顯示各大行業產業結構明顯差異，行業分類應為後續研究極為重要之因素。

表 4.1.1 各大類行業母體資料分布

	家數		從業員工人數		資產淨額		營業收入	
	(家)	分配比 (%)	(人)	分配比 (%)	(十億元)	分配比 (%)	(十億元)	分配比 (%)
總計	1,065,250	100.00	7,510,060	100.00	101,106.9	100.00	43,999.6	100.00
礦業及土石採取業	497	0.05	4,969	0.07	199	0.02	21.7	0.05
製造業	148,017	13.90	2,695,984	35.90	21,182.7	20.95	19,725.2	44.83
電力及燃氣供應業	166	0.02	32,147	0.43	1,679.0	1.66	515.3	1.17
用水供應及污染整治業	4,008	0.38	27,809	0.37	415.5	0.41	88.4	0.20
營造業	73,360	6.89	479,357	6.38	1,672.0	1.65	1,476.3	3.36
批發及零售業	490,017	46.00	1,889,924	25.17	5,788.5	5.73	11,496.5	26.13
運輸及倉儲業	19,512	1.83	320,700	4.27	3,207.3	3.17	1,072.6	2.44
住宿及餐飲業	88,739	8.33	323,692	4.31	532.4	0.53	384.2	0.87
資訊及通訊傳播業	11,907	1.12	179,371	2.39	1,524.3	1.51	753.5	1.71
金融及保險業	9,955	0.93	376,646	5.02	59,284.6	58.64	5,685.4	12.92
不動產業	15,675	1.47	83,932	1.12	2,284.1	2.26	586.0	1.33
專業、科學及技術服務業	40,719	3.82	179,356	2.39	357.1	0.35	373.7	0.85
支援服務業	21,056	1.98	262,978	3.50	289.5	0.29	231.6	0.53
公共行政及國防；強制性社會安全	4	0.00	4,594	0.06	1,233.9	1.22	705.3	1.60
教育服務業	14,428	1.35	80,980	1.08	58.8	0.06	56.8	0.13
醫療保健及社會工作服務業	27,548	2.59	349,956	4.66	996.5	0.99	567.4	1.29
藝術、娛樂及休閒服務業	17,919	1.68	68,613	0.91	291.9	0.29	87.2	0.20
其他服務業	81,723	7.67	149,052	1.98	288.9	0.29	172.5	0.39

貳、歷次普查母體結構分析

茲以 CHAID 產生樹狀圖分析 75 年至 95 年普查母體資料結構變化，各年皆以營業收入為準則變數，並挑選細類行業別、縣市鄉鎮、從業員工人數、從業員工薪資、總支出及資產淨額計 6 個特徵變數，在 $\alpha=0.05$ 的水準下，皆約分 20 層，其特徵變數留存結果如表 4.1.2 所示。

表 4.1.2 75 年至 95 年普查母體資料 CHAID 分析

年度	母體家數	層數	父節點/子節點	準則變數	特徵變數
75 年	611,161 家	21	40,000/20,000	營業收入	從業員工人數、從業員工薪資、總支出、縣市鄉鎮、資產淨額、細類行業別
80 年	704,236 家	23	40,000/20,000	營業收入	從業員工人數、從業員工薪資、總支出、縣市鄉鎮、資產淨額、細類行業別
85 年	819,296 家	19	50,000/25,000	營業收入	從業員工人數、從業員工薪資、總支出、縣市鄉鎮、資產淨額
90 年	888,659 家	20	50,000/25,000	營業收入	從業員工人數、從業員工薪資、總支出、資產淨額
95 年	1,065,250 家	17	60,000/30,000	營業收入	從業員工人數、從業員工薪資、總支出、縣市鄉鎮

普查母體家數自 75 年之 61 萬家，增至 95 年之 106 萬家，20 年來增加約 45 萬家或 74.3%，除家數增幅較大外，產業變遷亦相當快速，75 年及 80 年母體的細類行業別仍可明顯區分營業收入的多寡，惟至 85 年以後，細類行業別已無法明顯區分；另縣市鄉鎮除 90 年外，其餘年度皆可區分；資產淨額則除 95 年外，其餘年度亦可區分；至從業員工人數、從業員工薪資及總支出等 3 變數，20 年來均屬區分營收規模之重要因素。

第二節 樣本結構分析

壹、95 年普查樣本資料

95 年普查抽樣家數共計 72,982 家廠商，略述如次：

一、負責單位：約 4 萬 8 千餘家由各縣市政府主計室負責查填；餘約 2 萬 5 千家由分工機關負責查填。

二、調查表式：

(一) 抽樣調查甲表：查填廠商約 5 萬 7 千家，包含製造業、營造業、商業、運輸及倉儲業、金融及保險業、服務事業及補習教育及托育服務業等 7 類，均約包含 10 個問項，120 個細項，其中屬類別變數約 20 個，餘皆屬連續變數。

(二) 抽樣調查乙表：查填廠商約 1 萬 6 千家，包含製造業、營造業、商業、運輸及倉儲業、金融及保險業(一)、(二)、(三)及服務事業等 8 類，均約包含 14 個問項，180 個細項，其中屬類別變數約 20 個，餘皆屬連續變數。

三、分層方式：

(一) 全查層：含細類行業別家數小於 10 家者、醫院、診所、經濟部工業產銷存調查對象、金融業分工調查企業及其餘細類行業營業收入截略點以上之企業。

(二) CHAID 分層：含 195 個細行業別。

(三) OLAP Cubes 分層：含 275 個細行業別。

觀察樣本資料之家數、從業員工人數、營業收入及資產淨額等營運資料之行業分布如次(詳表 4.2.1)：

一、家數：95 年樣本家數以醫療保健及社會工作服務業約 2 萬 2 千家最多，占 30.37%；製造業 2 萬 1 千家次之，占 28.93%；強制性社會安全僅 4 家最少，僅占 0.01%。

二、從業員工人數：95年樣本從業員工人數約371萬人，以製造業近181萬人最多，占48.72%；批發及零售業約44萬人次之，占11.88%；礦業及土石採取業近3千人最少，僅占0.09%。

三、資產淨額：95年樣本資產淨額約92兆元，其中以金融及保險業58兆元居冠，所占比率達62.97%；製造業近20兆元次之，占21.41%；運輸及倉儲業近3兆元，占3.22%；教育服務業約50億元最少，僅占0.01%。

四、營業收入：95年樣本營業收入近34兆元，以製造業近18兆元最多，占52.68%；批發及零售業近6兆元次之，占17.31%；金融及保險業約5兆元，占16.14%；教育服務業近80億元最少，僅占0.02%。

表 4.2.1 各大類行業樣本資料分布

	家數		從業員工人數		資產淨額		營業收入	
	(家)	分配比(%)	(人)	分配比(%)	(十億元)	分配比(%)	(十億元)	分配比(%)
總計	72,982	100.00	3,705,744	100.00	92,228.6	100.00	33,853.0	100.00
礦業及土石採取業	169	0.23	3,285	0.09	16.4	0.02	15.9	0.05
製造業	21,116	28.93	1,805,473	48.72	19,747.3	21.41	17,832.2	52.68
電力及燃氣供應業	142	0.19	31,902	0.86	1,677.9	1.82	514.2	1.52
用水供應及污染整治業	405	0.55	14,126	0.38	392.7	0.43	61.1	0.18
營造業	3,053	4.18	111,177	3.00	1,069.3	1.16	699.2	2.07
批發及零售業	14,114	19.34	440,202	11.88	2,810.7	3.05	5,860.8	17.31
運輸及倉儲業	1,502	2.06	155,636	4.20	2,973.0	3.22	835.5	2.47
住宿及餐飲業	952	1.30	91,410	2.47	239.5	0.26	113.0	0.33
資訊及通訊傳播業	1,208	1.66	120,355	3.25	1,376.5	1.49	609.3	1.80
金融及保險業	1,183	1.62	342,816	9.25	58,078.4	62.97	5,463.3	16.14
不動產業	816	1.12	25,136	0.68	1,055.1	1.14	264.4	0.78
專業、科學及技術服務業	2,391	3.28	66,968	1.81	168.4	0.18	156.2	0.46
支援服務業	1,328	1.82	126,881	3.42	172.2	0.19	111.5	0.33
公共行政及國防；強制性社會安全	4	0.01	4,594	0.12	1,233.9	1.34	705.3	2.08
教育服務業	563	0.77	12,188	0.33	5.3	0.01	7.8	0.02
醫療保健及社會工作服務業	22,168	30.37	308,862	8.33	958.5	1.04	538.0	1.59
藝術、娛樂及休閒服務業	743	1.02	26,411	0.71	190.0	0.21	36.1	0.11
其他服務業	1,125	1.54	18,322	0.49	63.4	0.07	29.2	0.09

貳、列聯表分析

觀察大行業與地區別交叉之樣本家數分布情形，以北部地區占 46.5% 最多，南部地區 26.9% 次之，中部地區 24.5% 居第 3，金馬地區占 0.2% 最少；復進行二者之獨立性檢定，其 P -Value <0.05 ，Cramer'V=0.102 其 P -Value=0.000 <0.05 ，顯示就家數分布而言，大行業與地區別存在顯著差異。18 類大行業大致以北部地區為主，東部地區及金馬地區最少，如表 4.2.2 所示。除「強制性社會安全」外，觀察大類行業之樣本家數地區分布，除礦業及土石採取業以中部地區 36.7% 居首外，餘行業分布皆側重於北部地區，其中又以資訊及通訊傳播業、專業、科學及技術服務業、金融及保險業與運輸及倉儲業逾 6 成較為密集，用水供應及污染整治業最低，惟亦逾 4 成，而東部及金馬地區，各大類行業之分布最低，僅占總樣本家數之 2.2%。

表 4.2.2 大行業與地區別列聯表

	地區別					合計
	北部	中部	南部	東部	金馬	
礦業及土石採取業	51	62	20	36	0	169
	30.2%	36.7%	11.8%	21.3%	0.0%	100.0%
製造業	8830	6586	5541	144	15	21116
	41.8%	31.2%	26.2%	0.7%	0.1%	100.0%
電力及燃氣供應業	61	31	39	9	2	142
	43.0%	21.8%	27.5%	6.3%	1.4%	100.0%
用水供應及污染整治業	163	109	122	9	2	405
	40.2%	26.9%	30.1%	2.2%	0.5%	100.0%
營造業	1341	657	929	111	15	3053
	43.9%	21.5%	30.4%	3.6%	0.5%	100.0%
批發及零售業	7117	2697	3927	343	30	14114
	50.4%	19.1%	27.8%	2.4%	0.2%	100.0%
運輸及倉儲業	906	193	371	18	14	1502
	60.3%	12.8%	24.7%	1.2%	0.9%	100.0%
住宿及餐飲業	403	215	281	43	10	952
	42.3%	22.6%	29.5%	4.5%	1.1%	100.0%
資訊及通訊傳播業	942	102	151	10	3	1208
	78.0%	8.4%	12.5%	0.8%	0.2%	100.0%
金融及保險業	802	160	193	26	2	1183
	67.8%	13.5%	16.3%	2.2%	0.2%	100.0%
不動產業	480	145	185	6	0	816
	58.8%	17.8%	22.7%	0.7%	0.0%	100.0%

專業、科學及技術服務業	1638	350	370	31	2	2391
	68.5%	14.6%	15.5%	1.3%	0.1%	100.0%
支援服務業	729	235	320	39	5	1328
	54.9%	17.7%	24.1%	2.9%	0.4%	100.0%
公共行政及國防； 強制性社會安全	4	0	0	0	0	4
	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%
教育服務業	291	111	156	5	0	563
	51.7%	19.7%	27.7%	0.9%	0.0%	100.0%
醫療保健及社會工作服務業	9369	5767	6451	536	45	22168
	42.3%	26.0%	29.1%	2.4%	0.2%	100.0%
藝術、娛樂及休閒服務業	315	172	213	34	9	743
	42.4%	23.1%	28.7%	4.6%	1.2%	100.0%
其他服務業	480	272	338	29	6	1125
	42.7%	24.2%	30.0%	2.6%	0.5%	100.0%
合計	33922	17864	19607	1429	160	72982
	46.5%	24.5%	26.9%	2.0%	0.2%	100.0%

Cramer'V=0.102 ; P -Value=0.000

參、關聯性分析

95 年普查抽樣作業之 CHAID 分層部分，係以營業收入為準則變數，並考量從業員工人數、資產淨額及縣市鄉鎮之互動關係，以各最終節點(terminate nodes)間的離散程度最大為條件建構決策樹。茲以印刷業(行業代號 1611，計分 9 層，共 248 個樣本)與服裝及其配件批發業(行業代號 4552，計分 8 層，共 239 個樣本)為例，分析去除從業員工人數、資產淨額及縣市鄉鎮之互動關係後，各最終節點的樣本關聯程度。分析結果如次：

一、**兩連續變數**：自普查問項內容中挑選 18 個變數，於 $\alpha=0.05$ 之水準下，觀察相關係數(Pearson Correlation Coefficients)之顯著程度。茲就前揭二細類行業分述其結果如次：

(一)**印刷業**：先觀察營業收入與各變數間之相關係數，各層之營業收入皆與總支出之相關係數較高；且與中間消費、從業員工人數、勞動報酬、從業員工薪資存在高相關；而第 3 層、第 6 層、第 8 層與第 9 層與資產之相關性，則高於其他層別，如表 4.2.3.1 所示。另觀察各連續財務特徵值間之相關係數，各層從業員工薪資與勞動報酬皆存在高度相關性，而第 9 層之從業員工人

數、勞動報酬、從業員工薪資及資產淨額間亦皆顯著相關(詳附錄 1)。

表 4.2.3.1 印刷業營業收入與各連續特徵值之相關係數

層別	樣本數	變數名稱	相關係數	P-Value
第 1 層	7	總支出	0.97477	<.0001
		中間消費	0.99692	<.0001
		勞動報酬	0.99970	<.0001
		從業員工薪資	0.93383	0.0021
		從業員工人數	0.92527	0.0028
第 2 層	8	總支出	0.97477	<.0001
		中間消費	0.96026	0.0002
第 3 層	12	總支出	0.99938	<.0001
		勞動報酬	0.91664	<.0001
		中間消費	0.99814	<.0001
		從業員工薪資	0.85440	0.0004
		資產淨額	0.60828	0.0359
第 4 層	8	總支出	0.99702	<.0001
		勞動報酬	0.97031	<.0001
		中間消費	0.99436	<.0001
		從業員工薪資	0.90345	0.0021
第 5 層	15	總支出	0.99184	<.0001
		中間消費	0.93855	<.0001
		從業員工人數	0.75206	0.0012
		勞動報酬	0.68292	0.0050
		從業員工薪資	0.66986	0.0063
第 6 層	14	總支出	0.99650	<.0001
		中間消費	0.98554	<.0001
		勞動報酬	0.84920	0.0001
		從業員工薪資	0.82419	0.0003
		資產淨額	0.68947	0.0064
第 7 層	20	總支出	0.98631	<.0001
		中間消費	0.96978	<.0001
		存貨及存料資產	0.75086	0.0001
第 8 層	33	總支出	0.99341	<.0001
		從業員工人數	0.70635	<.0001
		從業員工薪資	0.69442	<.0001
		勞動報酬	0.68411	<.0001
		存貨及存料資產	0.64807	<.0001
		中間消費	0.99117	<.0001
		資產淨額	0.44165	0.0101
第 9 層	131	總支出	0.98897	<.0001
		勞動報酬	0.79159	<.0001
		從業員工薪資	0.78083	<.0001
		從業員工人數	0.75301	<.0001
		資產淨額	0.66444	<.0001
		中間消費	0.94715	<.0001

(二) 服裝及其配件批發業：先觀察營業收入與各變數間之相關係

數，各層之營業收入仍與總支出相關係數最高；第3層、第6層、第7層、第8層與資產之相關性，則高於其他層別；而僅第7層之營業收入與專業技術購入金額有關，如表4.2.3.2所示。另觀察各連續財務特徵值間之相關係數，各層從業員工薪資與勞動報酬皆存在高度相關性，僅第2層之從業員工人數與房屋及建築淨額之資產呈現負相關(詳附錄2)。

表 4.2.3.2 服裝及其配件批發業營業收入與各連續特徵值相關係數

層別	樣本數	變數名稱	相關係數	P-Value
第1層	7	總支出	0.98912	<.0001
第2層	15	總支出	0.99434	<.0001
		勞動報酬	0.92486	<.0001
		從業員工薪資	0.89107	<.0001
		從業員工人數	0.57232	0.0258
第3層	22	總支出	0.99861	<.0001
		中間消費	0.79892	<.0001
		勞動報酬	0.67186	0.0006
		從業員工薪資	0.67063	0.0006
		資產淨額	0.47203	0.0266
第4層	8	從業員工人數	0.46476	0.0293
		總支出	0.99702	<.0001
		勞動報酬	0.97031	<.0001
		中間消費	0.99436	<.0001
第5層	15	從業員工薪資	0.90345	0.0021
第6層	5	總支出	0.98683	<.0001
		總支出	0.99675	0.0002
		存貨及存料資產	0.93528	0.0196
第7層	42	資產淨額	0.92814	0.0229
		總支出	0.99109	<.0001
		中間消費	0.77352	<.0001
		勞動報酬	0.60585	<.0001
		從業員工薪資	0.62035	<.0001
		專業技術購入金額	0.43691	0.0038
第8層	65	資產淨額	0.38632	0.0115
		總支出	0.99794	<.0001
		中間消費	0.74598	<.0001
		資產淨額	0.42951	<.0001
		勞動報酬	0.28288	<.0001
		從業員工薪資	0.24804	<.0001

二、兩類別變數：仍以前揭二細類行業為例，自普查問項內容中挑選5個類別變數(變數定義如表4.2.3.3)，以Goodman & Kruskal(JASA, 1963)關聯測量法，計算消滅錯誤之比例及Lambda、Goodman and Kruskal Tau統計量，俾觀察兩兩變數間

之相關性。分析結果如次：

表 4.2.3.3 類別變數定義

變數名稱	變數內容
單位級別	1：獨立經營單位 8：有分支單位之企業
組織別	1：民營公司組織 2：民營獨資或合夥 3：民營其他組織 4：公營公司組織 5：公營非公司及其他組織
主要經營方式	1：製造 2：修配 3：代客加工 4：自行研發設計委外生產
使用電腦情形	1：有 2：無
使用人力派遣情形	1：有 2：無

(一) 印刷業：以「縣市別」預測各層類別變數時，皆以「組織別」之 Lambda 統計量最高，即可消滅的錯誤較多；第 7 層至第 9 層預測「單位級別」時，消滅錯誤皆很少；第 2 層預測「主要經營方式」時，消滅的錯誤較多，其餘各層消滅的錯誤則相對較少；第 1 層至第 4 層預測「使用電腦情形」時，消滅的錯誤較多，其餘各層消滅的錯誤則相對較少，如表 4.2.3.4 所示。

表 4.2.3.4 印刷業以縣市別預測各層類別變數之相關情形

層別	統計量	單位級別	組織別	主要經營方式	使用電腦情形
第 1 層	Lambda	-	1	-	1
	Tau	-	1	-	1
第 2 層	Lambda	-	0.750	0.667	0.750
	Tau	-	0.667	0.644	0.667
第 3 層	Lambda	-	0.750	0.000	0.500
	Tau	-	0.719	0.273	0.625
第 4 層	Lambda	-	1	-	1
	Tau	-	1	-	1
第 5 層	Lambda	-	0.714	0.000	0.000
	Tau	-	0.652	0.250	0.279
第 6 層	Lambda	-	1	0.000	-
	Tau	-	1	0.462	-
第 7 層	Lambda	0.000	0.500	0.000	-
	Tau	0.474	0.484	0.444	-
第 8 層	Lambda	0.000	0.750	0.000	0.000
	Tau	0.083	0.758	0.313	0.279
第 9 層	Lambda	0.000	0.125	0.000	0.000
	Tau	0.089	0.234	0.151	0.275

註：「-」指無法計算。

(二) 服裝及其配件服飾業：以「縣市別」預測各層類別變數時，「組織別」於第 1 層、第 4 層及第 5 層，「使用電腦情形」於第 1 層、第 4 層及第 6 層可完全消滅錯誤，Lambda 統計量均為 1；另「使用電腦情形」於第 2 層，「組織別」於第 2 層，「單位級別」於第 7 層可消滅的錯誤相對較多，如表 4.2.3.5 所示。

表 4.2.3.5 服裝及其配件批發業以縣市別預測各層類別變數之相關情形

層別	統計量	單位級別	組織別	使用電腦情形
第 1 層	Lambda	-	1	1
	Tau	-	0.500	1
第 2 層	Lambda	-	0.667	0.714
	Tau	-	0.676	0.688
第 3 層	Lambda	-	0.500	0.500
	Tau	-	0.588	0.511
第 4 層	Lambda	-	1	1
	Tau	-	1	1
第 5 層	Lambda	-	1	1
	Tau	-	1	1
第 6 層	Lambda	-	-	1
	Tau	-	-	1
第 7 層	Lambda	0.333	0.333	0.250
	Tau	0.491	0.392	0.417
第 8 層	Lambda	0.083	0.000	0.000
	Tau	0.173	0.323	0.175

註：「-」指無法計算。

三、連續變數與類別變數：仍以前揭二細類行業為例，將營業收入依 DH 分層法區分 20 層，並以 ETA(E)係數值衡量，ETA 的平方值 E^2 表示消滅誤差的比率，分別觀察 5 個類別變數與營業收入相關程度。分析結果如次：

(一) 印刷業：各層均以「縣市別」可消滅錯誤最多，其中以第 4 層可完全消滅錯誤最多，第 1 層及第 9 層消滅錯誤 22.66% 最少 ($E^2=0.2266$)；另各層均以「單位級別」可消滅錯誤最少，皆在 0.03% 以下，如表 4.2.3.6 所示。

(二) 服裝及其配件批發業：各層均以「縣市別」可消滅錯誤最多，其中以第 5 層、第 6 層可完全消滅錯誤最多，第 7 層消滅錯誤 29.27% 最少 ($E^2=0.2927$)；另各層均以「單位級別」可消滅錯誤最少，皆在 0.03% 以下，如表 4.2.3.7 所示。

表 4.2.3.6 印刷業營業收入與各類別變數之相關程度

層別	變數名稱	ETA 係數(E)	ETA 係數平方(E ²)
第 1 層	單位級別	0.039	0.0015
	組織別	0.266	0.0708
	主要經營方式	0.298	0.0888
	使用電腦情形	0.264	0.0697
	縣市別	0.476	0.2266
第 2 層	單位級別	-	-
	組織別	0.222	0.0493
	主要經營方式	0.078	0.0061
	使用電腦情形	0.556	0.3091
	縣市別	0.861	0.7413
第 3 層	單位級別	-	-
	組織別	0.302	0.0912
	主要經營方式	-	-
	使用電腦情形	0.302	0.0912
	縣市別	0.718	0.5155
第 4 層	單位級別	-	-
	組織別	0.626	0.3919
	主要經營方式	-	-
	使用電腦情形	0.120	0.0144
	縣市別	1	1
第 5 層	單位級別	-	-
	組織別	0.056	0.0031
	主要經營方式	0.331	0.1096
	使用電腦情形	0.414	0.1714
	縣市別	0.742	0.5506
第 6 層	單位級別	-	-
	組織別	0.053	0.0028
	主要經營方式	0.320	0.1024
	使用電腦情形	-	-
	縣市別	0.498	0.2480
第 7 層	單位級別	0.140	0.0196
	組織別	0.210	0.0441
	主要經營方式	-	-
	使用電腦情形	0.204	0.0416
	縣市別	0.499	0.2490
第 8 層	單位級別	0.146	0.0213
	組織別	0.283	0.0801
	主要經營方式	0.375	0.1406
	使用電腦情形	0.149	0.0222
	縣市別	0.721	0.5198
第 9 層	單位級別	0.039	0.0015
	組織別	0.266	0.0708
	主要經營方式	0.036	0.0013
	使用電腦情形	0.264	0.0697
	縣市別	0.476	0.2266

註：「-」指無法計算。

表 4.2.3.7 服裝及其配件批發業營業收入與各類別變數之相關程度

層別	變數名稱	ETA 係數(E)	ETA 係數平方(E ²)
第 1 層	單位級別	-	-
	組織別	0.714	0.5098
	使用電腦情形	0.163	0.0266
	縣市別	0.994	0.9880
第 2 層	單位級別	-	-
	組織別	0.053	0.0028
	使用電腦情形	0.498	0.2480
	縣市別	0.642	0.4122
第 3 層	單位級別	-	-
	組織別	0.123	0.0151
	使用電腦情形	0.316	0.0999
	縣市別	0.717	0.5141
第 4 層	單位級別	-	-
	組織別	0.307	0.0942
	使用電腦情形	0.662	0.4382
	縣市別	0.804	0.6464
第 5 層	單位級別	-	-
	組織別	0.456	0.2079
	使用電腦情形	0.559	0.3125
	縣市別	1	1
第 6 層	單位級別	-	-
	組織別	-	-
	使用電腦情形	0.134	0.0180
	縣市別	1	1
第 7 層	單位級別	0.133	0.0177
	組織別	0.133	0.0177
	使用電腦情形	0.440	0.1936
	縣市別	0.541	0.2927
第 8 層	單位級別	0.146	0.0213
	組織別	0.137	0.0188
	使用電腦情形	0.192	0.0369
	縣市別	0.552	0.3047

註：「-」指無法計算

第三節 95 年普查母體與樣本資料分析

壹、抽出率

95 年普查母體計 1,065,250 家，共抽取樣本 72,982 家，總抽出率為 6.85%，依各大類行業分別觀察樣本抽出率分布如次：

- 一、批發及零售業母體家數 49 萬家最多，占母體家數之 46.00%；樣本分配比 19.34%，抽出率 2.88%。
- 二、製造業母體家數近 15 萬家次之，占母體家數之 13.90%；樣本分配比 28.93%，抽出率 14.27%。
- 三、住宿及餐飲業母體家數近 9 萬家，占母體家數之 8.33%；樣本分配比 1.30%，抽出率 1.07%。
- 四、其他服務業母體家數逾 8 萬家，占母體家數之 7.67%；樣本分配比 1.54%，抽出率 1.38%。
- 五、營造業母體家數逾 7 萬家，占母體家數之 6.89%；樣本分配比 4.18%，抽出率 4.16%。

其餘各大業之母體家數分配比皆小於 5.00%，如表 4.3.1 所示。

普查辦理期間，為減輕廠商填表負荷，爰協調各機關，將部分常川調查與普查合併辦理或停辦，如醫療機構專案調查、工廠校正暨營運調查、工商企業經營概況調查(按年)、營造業經濟概況調查及金融業經營概況調查等；普查亦配合前揭調查需要，將醫療保健及社會工作服務業、製造業與金融及保險業之抽出率分別提升為 80.47%、14.27%及 11.88%。

另礦業及土石採取業、電力及燃氣供應業、用水供應及污染整治業、資訊及通訊傳播業與公共行政及國防(強制性社會安全)因行業特殊家數較少，抽出率皆逾 10%；各大類行業中，以住宿及餐飲業抽出率僅 1.07%最低。

表 4.3.1 各大類行業抽出率分布

	母體家數		樣本家數		抽出率 (C)=(B)÷(A) ×100%
	家數(A)	分配比(%)	家數(B)	分配比(%)	
總計(總平均)	1,065,250	100.00	72,982	100.00	6.85
礦業及土石採取業	497	0.05	169	0.23	34.00
製造業	148,017	13.90	21,116	28.93	14.27
電力及燃氣供應業	166	0.02	142	0.19	85.54
用水供應及污染整治業	4,008	0.38	405	0.55	10.10
營造業	73,360	6.89	3,053	4.18	4.16
批發及零售業	490,017	46.00	14,114	19.34	2.88
運輸及倉儲業	19,512	1.83	1,502	2.06	7.70
住宿及餐飲業	88,739	8.33	952	1.30	1.07
資訊及通訊傳播業	11,907	1.12	1,208	1.66	10.15
金融及保險業	9,955	0.93	1,183	1.62	11.88
不動產業	15,675	1.47	816	1.12	5.21
專業、科學及技術服務業	40,719	3.82	2,391	3.28	5.87
支援服務業	21,056	1.98	1,328	1.82	6.31
公共行政及國防 (強制性社會安全)	4	0.00	4	0.01	100.00
教育服務業	14,428	1.35	563	0.77	3.90
醫療保健及社會工作服 務業	27,548	2.59	22,168	30.37	80.47
藝術、娛樂及休閒服務業	17,919	1.68	743	1.02	4.15
其他服務業	81,723	7.67	1,125	1.54	1.38

貳、層數分析

95 年普查除全查樣本外，餘按各細類行業家數，分採 OLAP Cubes 或 CHAID 分層，母體分層共計 3,681 層，樣本計 3,352 層(均含全查層)，樣本實際回表資料較母體數少 329 層，空層比率達 8.94%。

依大類行業別觀察，製造業與批發及零售業細類行業分層數較多，空層數分別為 150 層及 53 層；製造業空層比率 9.49%，批發及零售業 6.45%；各大類行業空層比率以藝術、娛樂及休閒服務業 21.05% 最高，營造業 1.83% 最低，如表 4.3.2 所示。

表 4.3.2 各大類行業層數分布

	細類行業數		分層數			
	母體	樣本	母體 (A)	樣本 (B)	空層數 (C)=(A)-(B)	空層比率 (C)÷(A) ×100%
總計(總平均)	495	494	3,681	3,352	329	8.94
礦業及土石採取業	3	3	19	17	2	10.53
製造業	212	211	1,580	1,430	150	9.49
電力及燃氣供應業	3	3	11	11	-	-
用水供應及污染整治業	8	8	54	48	6	11.11
營造業	11	11	109	107	2	1.83
批發及零售業	89	89	822	769	53	6.45
運輸及倉儲業	26	26	157	141	16	10.19
住宿及餐飲業	6	6	58	46	12	20.69
資訊及通訊傳播業	21	21	152	132	20	13.16
金融及保險業	26	26	113	105	8	7.08
不動產業	5	5	45	42	3	6.67
專業、科學及技術服務業	20	20	156	147	9	5.77
支援服務業	26	26	170	148	22	12.94
公共行政及國防 (強制性社會安全)	1	1	1	1	-	-
教育服務業	5	5	22	20	2	9.09
醫療保健及社會工作服務業	11	11	47	45	2	4.26
藝術、娛樂及休閒服務業	11	11	76	60	16	21.05
其他服務業	11	11	89	83	6	6.74

第四節 樣本精確度分析

精確度(precision)係指估計值精確程度，通常以重複試驗結果與母數接近之程度衡量。95 年普查抽樣作業 CHAID 分層係以營業收入為準則變數，從業員工人數、資產淨額及地區別為特徵變數；OLAP Cubes 則以營業收入、從業員工人數分層，爰將各細類行業別扣除全查層資料後，分別計算樣本與母體之營業收入、從業員工人數及資產淨額等 3 項企業營運之重要指標，據以衡量樣本之精確程度。

壹、營業收入

先將營業收入取 0.134912 次方轉換，再以截尾平均數方法，比較 461 個細類行業母體及樣本每企業營業收入之平均數差異情形，其中以家用洗衣設備製造業(2853)、布疋零售業(4731)及電視傳播業(6021)差異最小；而汽車零售業(4841)、其他綜合商品零售業(4719)與服裝及其配件零售業(4732)差異最大。差距比例(R)介於 2~3 倍之細類行業個數最多，占 32.10%，以製造業 68 個細類行業居多；差距比例(R)9 倍以上的細類行業占 2.17%，以批發及零售業 8 個細類行業居多，如表 4.4.1 所示。以印刷業(1611)為例，其母體及樣本資料之分布及轉換後分布分別詳附錄 3 至附錄 6。

表 4.4.1 母體與樣本平均每企業營業收入

$R = \frac{\text{樣本每企業營業收入}}{\text{母體每企業營業收入}}$	細業別個數 (%)	累積百分比 (%)
$R < 1$	17 (3.69%)	3.69%
$1 \leq R < 2$	83 (18.00%)	21.69%
$2 \leq R < 3$	148 (32.10%)	53.80%
$3 \leq R < 4$	103 (22.34%)	76.14%
$4 \leq R < 5$	50 (10.85%)	86.98%
$5 \leq R < 6$	27 (5.86%)	92.84%
$6 \leq R < 7$	14 (3.04%)	95.88%
$7 \leq R < 8$	8 (1.74%)	97.61%
$8 \leq R < 9$	1 (0.22%)	97.83%
$R \geq 9$	10 (2.17%)	100.00%
合計	461 (100.00%)	

貳、從業員工人數

先將從業員工人數取 0.134912 次方轉換，再以截尾平均數方法，比較 461 個細類行業母體及樣本每企業從業員工人數之平均數差異情形，以護理照顧服務業(8701)、電視傳播業(6021)及家用洗衣設備製造業(2853)差異最小；而服裝及其配件零售業(4732)、其他綜合商品零售業(4719)及汽車零售業(4841)差異最大；差距比例(R)介於 2~3 倍之細類行業個數最多，占 45.99%，以製造業 113 個細類行業居多；相差 9 倍以上的細類行業占 0.65%，皆屬批發及零售業，如表 4.4.2 所示。

表 4.4.2 母體與樣本平均每企業從業員工人數

$R = \frac{\text{樣本每企業從業員工}}{\text{母體每企業從業員工}}$	細業別個數 (%)	累積百分比 (%)
$R < 1$	10 (2.17%)	2.17%
$1 \leq R < 2$	129 (27.98%)	30.15%
$2 \leq R < 3$	212 (45.99%)	76.14%
$3 \leq R < 4$	77 (16.70%)	92.84%
$4 \leq R < 5$	20 (4.34%)	97.18%
$5 \leq R < 6$	9 (1.95%)	99.13%
$6 \leq R < 7$	1 (0.22%)	99.35%
$7 \leq R < 8$	0 (0.00%)	99.35%
$8 \leq R < 9$	0 (0.00%)	99.35%
$R \geq 9$	3 (0.65%)	100.00%
合計	461 (100.00%)	

參、資產淨額

先將資產淨額取 0.134912 次方轉換，再以截尾平均數方法，比較 461 個細類行業母體及樣本每企業資產淨額之平均數差異情形，以洗衣業(9610)、通訊傳播設備修理業(9523)及理髮及美容業(9620)差異最小；而其他綜合商品零售業(4719)、汽車零售業(4841)與運動及娛樂用品租賃業(7721)差異最大；差距比例(R)介於 2~3 倍之細類行業個數最多，占 20.17%，以製造業 50 個細類行業居多；相差 9 倍以上的細類行業占 4.12%，以批發及零售業 10 個細類行業居多，如表 4.4.3 所示。

表 4.4.3 母體與樣本平均每企業資產淨額

$R = \frac{\text{樣本每企業資產淨額}}{\text{母體每企業資產淨額}}$	細業別個數 (%)	累積百分比 (%)
$R < 1$	42 (9.11%)	9.11%
$1 \leq R < 2$	82 (17.79%)	26.90%
$2 \leq R < 3$	93 (20.17%)	47.07%
$3 \leq R < 4$	89 (19.31%)	66.38%
$4 \leq R < 5$	56 (12.15%)	78.52%
$5 \leq R < 6$	36 (7.81%)	86.33%
$6 \leq R < 7$	19 (4.12%)	90.46%
$7 \leq R < 8$	18 (3.90%)	94.36%
$8 \leq R < 9$	7 (1.52%)	95.88%
$R \geq 9$	19 (4.12%)	100.00%
合計	461 (100.00%)	

綜合前述，樣本精確度實與各細類行業之樣本抽出率存在密切關係，如母體及樣本特徵值截尾平均數差異較小之電視傳播業(6021)及家用洗衣設備製造業(2853)，樣本抽出率分別為高達 96.49%及 47.83%，故精確度較高；至其他綜合商品零售業(4719)及汽車零售業(4841)，樣本抽出率僅 4.84%及 4.56%，精確度則顯著低於其他細類行業。

第五節 行業別的分類（以製造業為例）

茲以製造業樣本為例進行 CART 分析，並以營業收入為依變項，各項企業營運資訊為自變項，透過分割技術，將製造業樣本資料以二元及序列展開方式進行分類，進而挑選影響營業收入最重要的自變項。分析流程如下：

- 一、轉換依變數：觀察樣本之營業收入分布，資料明顯呈現右偏(圖 4.5.1)，為了解各變項的重要性，並分析變項間的離散情形，爰將營業收入取 $\ln(\text{營業收入}+1)$ 轉換，使 $\ln(\text{營業收入}+1)$ 資料分布近似對稱(偏態係數為 2.785，如圖 4.5.2)，並以其為依變項進行分析。

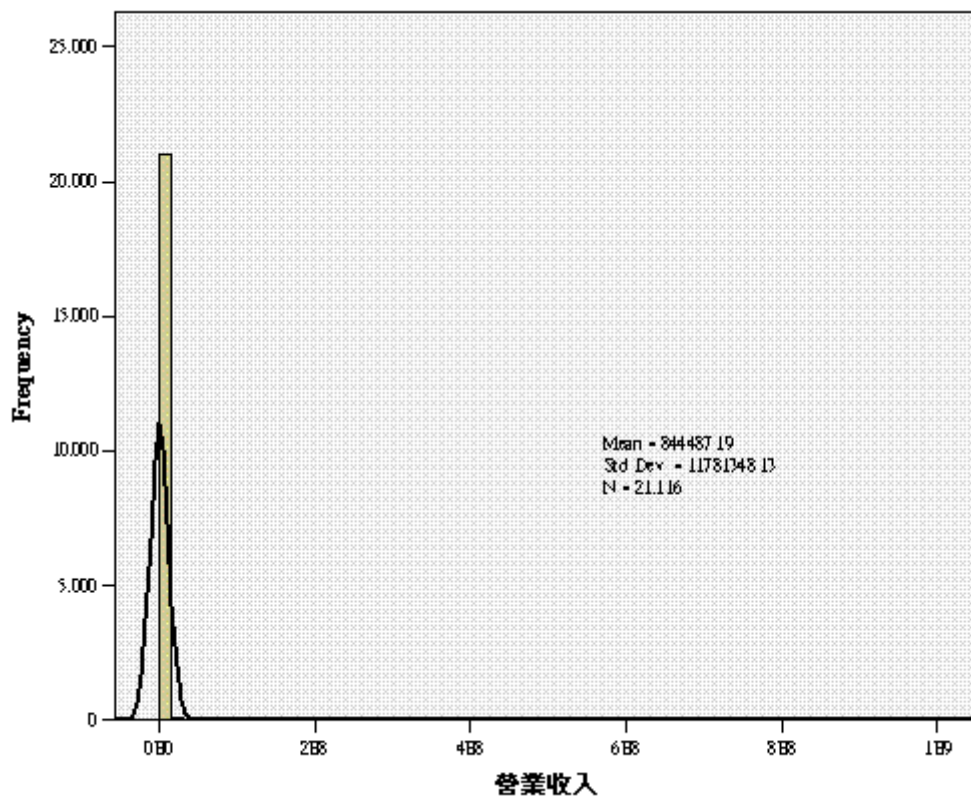


圖 4.5.1 製造業樣本資料營業收入分布

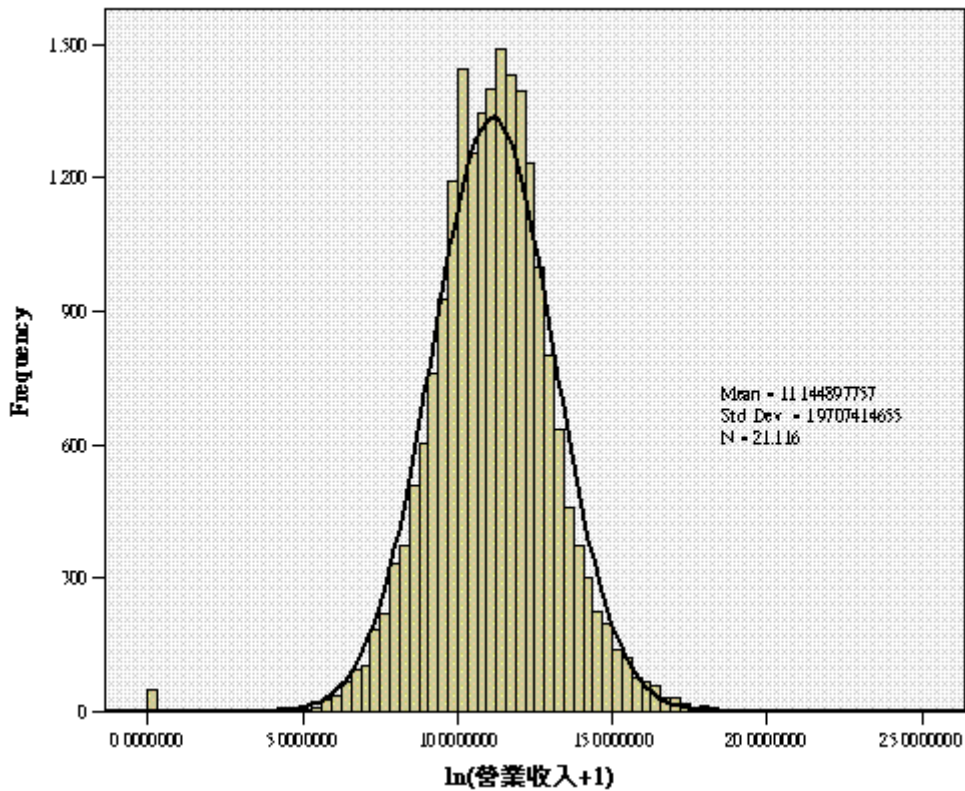


圖 4.5.2 製造業樣本資料 $\ln(\text{營業收入}+1)$ 分布

二、**挑選自變數**：為了解不同自變項之組成對依變項的影響，爰自普查問項內容挑選 26 個自變項，其中 3 個變項為類別變項，餘皆為連續變項，如表 4.5.1 所示。

表 4.5.1 CART 分析變數屬性說明－製造業

變數型態	變數名稱	變數內容
類別變數	細類行業別	中華民國行業標準分類第 8 次修訂
	主要經營方式	1：製造 2：修配 3：代客加工 4：自行研發設計委外生產
	組織別	1：民營公司組織 2：民營獨資或合夥 3：民營其他組織 4：公營公司組織 5：公營非公司及其他組織

變數型態	變數名稱	變數內容
連續變數	土地面積	平方公尺
	樓地板面積	平方公尺
	從業員工人數	
	從業員工薪資	千元
	中間消費	千元
	平均每家使用派遣勞工人數	
	平均每家經營派遣人數	
	租金收入	千元
	利息收入	千元
	資產－存貨及存料	千元
	資產－土地	千元
	資產－房屋及建築	千元
	資產－運輸設備資產	千元
	資產－機械及什項設備	千元
	資產－未完工程及預付購置設備	千元
	資產－流動資產	千元
	資產－資產淨額	千元
	無形投資－研究發展支出	千元
	無形投資－員工訓練支出	千元
	無形投資－市場行銷支出	千元
	無形投資－購買電腦軟體資料庫支出	千元
無形投資－專業技術購入	千元	
無形投資－專業技術銷售	千元	

三、產生 CART 樹狀圖(附錄 7)：在節點 16 時， $N(16)=479$ ，

$$N(16_L) = 128, \quad N(16_R) = 351, \quad N = 21116, \quad \bar{y}(16_L) = 5.9114,$$

$$\bar{y}(16_R) = 7.2186, \quad \bar{y} = 11.1449, \quad SST = \sum_i^N (y_i - \bar{y})^2 = 82,011, \quad \text{可知}$$

$$\hat{C}_{x_i}(16) = \frac{\frac{128}{479}(5.9114 - 11.1449)^2 + \frac{351}{479}(7.2186 - 11.1449)^2}{\frac{82001}{21116}} = 4.793079$$

最大，爰初步選取中間消費及從業員工薪資兩變數，如表 4.5.2 所示。

表 4.5.2 CART 分析變數初步選取－製造業

節點	變數組成	\hat{C}_{x_i}
16	中間消費、從業員工薪資	4.793079
17	中間消費、從業員工薪資	2.094170
18	中間消費、從業員工薪資、主要經營方式	0.929969
19	中間消費、主要經營方式、流動資產	0.161516
20	中間消費、主要經營方式	0.475615
22	中間消費、主要經營方式	0.061602
23	中間消費、主要經營方式、流動資產	0.028986
24	中間消費、從業員工薪資、主要經營方式	0.165924
25	中間消費、流動資產	0.486371
26	中間消費、流動資產	1.159218
28	中間消費、流動資產	2.690014
29	中間消費、流動資產	4.485556

四、計算 I_{x_i} 及 $R^2 x_i$ ：以中間消費及從業員工薪資重新進行 CART 分析，當移除其中一變項 X_i ，即計算 I_{x_i} 及 $R^2 x_i$ 。發現當移除中間消費時（CART 樹狀圖如附錄 8）， I_{x_i} 及 $R^2 x_i$ 皆較小（表 4.5.3），故對營業收入而言，中間消費為最重要的變項。

五、分層變數選取：因母體資料檔僅含財務報表大分類總計資料，故於分層變數選取時，將綜合考量各副母體內與中間消費及從業員工薪資相關程度較高，且涵括於母體資料檔之變項，以提升分層效果。

表 4.5.3 I_{x_i} 及 $R^2 x_i$ 值－製造業

移除變數	I_{x_i}	$R^2 x_i$
中間消費	1.4525	0.823351
從業員工薪資	1.5431	0.896326

第六節 交叉表中變項組合之次數分析

對數線性模型(Log-Linear Model)適用於分析列聯表變數間各水準組合之觀察次數，該模型係將某些變數(如營業收入、員工人數)之自然順序訊息予以忽略，並經由轉換後，視為名義變數(Nominal)；由變異數分析法，了解各變數間的效應及互動。將 95 年普查之樣本企業總數，藉各類變數形成之列聯表資料，可表示企業配置次數分布，惟企業次數受何變數影響較大，則須針對變數間之交互作用予以剖析。95 年普查運用於抽樣設計之變數計有營業收入、地區別及行業別等，其中營業收入分布因呈現極度偏斜，爰先加 1(為維持正值)後取對數，再依 DH 分層法分為 5 層，如表 4.6.1 所示。

表 4.6.1 95 年普查抽樣資料營業收入(y)DH 分層

$\ln(y+1)$	$f(\ln(y+1))$	$\sqrt{f(\ln(y+1))}$	$cum\sqrt{f(\ln(y+1))}$
未及 1	245	15.65	15.65
1~未及 2	11	3.32	18.97
2~未及 3	26	5.10	24.07
3~未及 4	48	6.93	31.00
4~未及 5	136	11.66	42.66
5~未及 6	800	28.28	70.94
6~未及 7	3,161	56.22	127.16
7~未及 8	7,234	85.05	212.21
8~未及 9	14,027	118.43	330.64
9~未及 10	14,086	118.68	449.32
10~未及 11	10,641	103.16	552.48
11~未及 12	9,415	97.03	649.51
12~未及 13	6,937	83.29	732.80
13~未及 14	3,516	59.29	792.09
14~未及 15	1,539	39.23	831.32
15~未及 16	669	25.87	857.19
16~未及 17	310	17.61	874.80
17~未及 18	118	10.86	885.66
18~未及 19	38	6.16	891.82
19~未及 20	21	4.58	896.40
20 以上	4	2.00	898.40
合計	72,982	898.40	分 5 組(切點：179.68)

復將表 4.6.1 中轉換並分為 5 組後之營業收入，與地區別 4 組(1=

北部，2=中部，3=南部；4=東部及金馬)及行業別 6 組(1=製造業、礦業及土石採取業；2=營造業；3=水電燃氣業(公用事業)、污染防治業、運輸及倉儲業；4=批發及零售業、住宿及餐飲業；5=金融及保險業、不動產及租賃業；6=服務事業)交叉組成(地區×營業收入×行業別)之 4×5×6 列聯表 (詳附錄 9)。茲以地區別為第 1 變數效應(α_i)，營業收入為第 2 變數效應(η_j)，行業別為第 3 變數效應(λ_k)，當地區、營業收入與行業別任二效應間，及三者間之交互作用皆呈顯著時，交互作用其取對數之家數分布 ($Ln(f_{ijk})$) 可由以下飽和對數線性模型 (Saturated Loglinear Model) 模型描述：

飽和對數線性模型(Saturated Loglinear Model)

$$Ln(f_{ijk}) = \mu_{..} + \alpha_i + \eta_j + \lambda_k + (\alpha\eta)_{ij} + (\alpha\lambda)_{ik} + (\eta\lambda)_{jk} + (\alpha\eta\lambda)_{ijk} + \varepsilon_{ijk}$$

如將飽和模型任一交互作用效應移除，即成非飽和模型 (Non-saturated Loglinear Model)。由表 4.6.2、表 4.6.3 及表 4.6.4 之最大概似變異數分析(Maximum Likelihood ANOVA)結果，可知模型各效應之 P -Value 皆小於顯著水準，即地區、營業收入與行業效應及其交互作用皆呈顯著，但其模型檢定結果 P -Value < 0.05，故儘管前述效應存在某種程度關係，但此關係顯然不夠強烈到足以構成一個完整模型的程度。爰採最大資訊量估計法，觀察企業家數受何變項影響程度較大。

表 4.6.2 最大概似變異數分析—飽和模型

Source	DF	Chi-Square	Pr > ChiSq
地區	3	4734.78	<.0001
營業收入	4	1773.06	<.0001
行業別	5	3546.88	<.0001
地區×營業收入	12	1096.12	<.0001
地區×行業別	15	1147.30	<.0001

營業收入×行業別	20	3249.85	<.0001
地區×營業收入×行業別	60	237.98	<.0001
Likelihood Ratio	0	.	.

表 4.6.3 最大概似變異數分析—非飽和模型(移除三個效應之交互作用效應項)

Source	DF	Chi-Square	Pr > ChiSq
地區	3	6175.29	<.0001
營業收入	4	2950.93	<.0001
行業別	5	6682.16	<.0001
地區×營業收入	12	2562.31	<.0001
地區×行業別	15	1957.63	<.0001
營業收入×行業別	20	15135.26	<.0001
Likelihood Ratio	60	243.12	<.0001

表 4.6.4 最大概似變異數分析—非飽和模型(移除所有交互作用效應)

Source	DF	Chi-Square	Pr > ChiSq
地區	3	18087.71	<.0001
營業收入	4	11659.48	<.0001
行業別	5	37090.51	<.0001
Likelihood Ratio	107	26303.16	<.0001

由附錄 10 之 Estimate 欄找到對數線性模型參數估計值(α_i 、 η_j 、 λ_k)，可進行如下之各變項效應之最大資訊量估計。其中樣本資料地區效應之最大資訊量估計值可由(c)式求得：

$$\begin{aligned}
 R \hat{\alpha}_i &= \max_{\text{固定}j,k} \left(\overbrace{\mu_{..} + \alpha_i + (\alpha\eta)_{i/j} + (\alpha\lambda)_{i/k} + (\alpha\eta\lambda)_{i/jk}} \right) \\
 &\quad - \min_{\text{固定}j,k} \left(\overbrace{\mu_{..} + \alpha_i + (\alpha\eta)_{i/j} + (\alpha\lambda)_{i/k} + (\alpha\eta\lambda)_{i/jk}} \right) \\
 &= \left(\max(\hat{\alpha}_i) - \min(\hat{\alpha}_i) \right) + \left(\max_{\text{固定}j} \left(\overbrace{(\alpha\eta)_{i/j}} \right) - \min_{\text{固定}j} \left(\overbrace{(\alpha\eta)_{i/j}} \right) \right) \\
 &+ \left(\max_{\text{固定}k} \left(\overbrace{(\alpha\lambda)_{i/k}} \right) - \min_{\text{固定}k} \left(\overbrace{(\alpha\lambda)_{i/k}} \right) \right) + \left(\max_{\text{固定}jk} \left(\overbrace{(\alpha\eta\lambda)_{i/jk}} \right) - \min_{\text{固定}jk} \left(\overbrace{(\alpha\eta\lambda)_{i/jk}} \right) \right) \dots\dots (c)
 \end{aligned}$$

固定 j, k	樣本 $R \hat{\alpha}_i$
---------	-----------------------

營業收入 $j=1$	$(3.758-1.781) + (1.114+0.052) + (2.757-1.892) + (1.718-1.571)$	=4.155
行業別 $k=1$		

而樣本資料營業收入效應之最大資訊量估計值可由(d)式求得

$$\begin{aligned}
R \hat{\eta}_j &= \max_{\text{固定}i,k} (\overbrace{\mu_{..} + \eta_j + (\alpha\eta)_{j/i} + (\eta\lambda)_{j/k} + (\alpha\eta\lambda)_{j/ik}}) \\
&\quad - \min_{\text{固定}i,k} (\overbrace{\mu_{..} + \eta_j + (\alpha\eta)_{j/i} + (\eta\lambda)_{j/k} + (\alpha\eta\lambda)_{j/ik}}) \\
&= (\max(\hat{\eta}_i) - \min(\hat{\eta}_i)) + (\max_{\text{固定}i} (\overbrace{(\alpha\eta)_{j/i}}) - \min_{\text{固定}i} (\overbrace{(\alpha\eta)_{j/i}})) \\
&+ (\max_{\text{固定}k} (\overbrace{(\eta\lambda)_{j/k}}) - \min_{\text{固定}k} (\overbrace{(\eta\lambda)_{j/k}})) + (\max_{\text{固定}ik} (\overbrace{(\alpha\eta\lambda)_{j/ik}}) - \min_{\text{固定}ik} (\overbrace{(\alpha\eta\lambda)_{j/ik}})) \dots\dots (d)
\end{aligned}$$

固定 i, k	樣本 $R \hat{\eta}_j$	
地區 $i=1$	$(3.559-1.099) + (1.114+0.035) + (2.586+0.990) + (2.002-1.412)$	=7.775
行業別 $k=1$		

最後，樣本資料行業別效應之最大資訊量估計值可由(e)式得

$$\begin{aligned}
R \hat{\lambda}_k &= \max_{\text{固定}i,j} (\overbrace{\mu_{..} + \lambda_k + (\alpha\lambda)_{k/i} + (\eta\lambda)_{k/j} + (\alpha\eta\lambda)_{k/ij}}) \\
&\quad - \min_{\text{固定}i,j} (\overbrace{\mu_{..} + \lambda_k + (\alpha\lambda)_{k/i} + (\eta\lambda)_{k/j} + (\alpha\eta\lambda)_{k/ij}}) \\
&= (\max(\hat{\lambda}_k) - \min(\hat{\lambda}_k)) + (\max_{\text{固定}i} (\overbrace{(\alpha\lambda)_{k/i}}) - \min_{\text{固定}i} (\overbrace{(\alpha\lambda)_{k/i}})) \\
&+ (\max_{\text{固定}j} (\overbrace{(\eta\lambda)_{k/j}}) - \min_{\text{固定}j} (\overbrace{(\eta\lambda)_{k/j}})) + (\max_{\text{固定}ij} (\overbrace{(\alpha\eta\lambda)_{k/ij}}) - \min_{\text{固定}ij} (\overbrace{(\alpha\eta\lambda)_{k/ij}})) \dots\dots (e)
\end{aligned}$$

固定 i, j	樣本 $R \hat{\lambda}_k$	
地區 $i=1$	$(1.099-0.143) + (1.892-0.110) + (3.455-0.473) + (2.411-0.035)$	=8.096
營業收入 $j=1$		

$$\text{由 } \hat{S}_{\hat{\alpha}_i} = \frac{R\hat{\alpha}_i}{R\hat{\alpha}_i + R\hat{\eta}_j + R\hat{\lambda}_k} \quad , \quad \hat{S}_{\hat{\eta}_j} = \frac{R\hat{\eta}_j}{R\hat{\alpha}_i + R\hat{\eta}_j + R\hat{\lambda}_k} \quad ,$$

$\hat{S}_{\hat{\lambda}_k} = \frac{R\hat{\lambda}_k}{R\hat{\alpha}_i + R\hat{\eta}_j + R\hat{\lambda}_k}$ 可估出樣本資料地區、營業收入、行業別各效應

能解釋整體模型效應之比率為 $\hat{S}_{\hat{\alpha}_i}$ 、 $\hat{S}_{\hat{\eta}_j}$ 、 $\hat{S}_{\hat{\lambda}_k}$ 。其值為 0.21、0.39、0.40，

即企業家數分布解釋比率最大之變項是行業，其次是營業收入及地區。

第七節 交叉表中變項類別之特定對應組合關聯分析

企業家數分布重要變項（地區、營業收入、行業）各效應能解釋整體模型效應之比率雖已算出，但變項類別之特定對應組合之關聯強度似乎更重要，因為這有助於了解現行樣本配置對企業家數分布之影響。當行業固定為製造業、礦業及土石採取業時，原 3 維列聯表退化成 2 維列聯表如表 4.7.1。

表 4.7.1 列聯表一固定製造業、礦業及土石採取業

地區	營收					Active Margin
	低營收	中低營收	中營收	中高營收	高營收	
北部	259	365	2645	3764	1848	8881
中部	406	592	2579	2464	607	6648
南部	347	527	2103	1919	665	5561
東部金馬	15	24	98	52	6	195
Active Margin	1027	1508	7425	8199	3126	21285

利用表 4.7.2 的 inertia 與表 4.7.3、表 4.7.4 的 score in dimension 求得某兩個類別之間的相關強度：

北部企業與高營業收入相關測度值＝

$$\begin{aligned} & \sqrt{\dim_1 \text{Inertia}} \times (\text{var}_1 \text{ score in dim}_1) \times (\text{var}_2 \text{ score in dim}_1) + \\ & \sqrt{\dim_2 \text{Inertia}} \times (\text{var}_1 \text{ score in dim}_2) \times (\text{var}_2 \text{ score in dim}_2) = \\ & \sqrt{0.046} \times (0.767) \times (0.677) + \sqrt{0.001} \times (-0.034) \times (0.226) = 0.1111 \end{aligned}$$

餘相關測度值詳表 4.7.5，故其相關都呈現正相關，其中以北部企業與高營業收入之關係最強，以東部及金馬地區企業與低營業收入之關係最弱。

表 4.7.2 摘要表—固定製造業、礦業及土石採取業

Summary

Dimension	Singular Value	Inertia	Proportion of Inertia		Confidence Singular Value	
			Accounted for	Cumulative	Standard Deviation	Correlation
1	.215	.046	.975	.975	.002	-.183
2	.034	.001	.025	1.000	.001	
3	.002	.000	.000	1.000		
Total		.048	1.000	1.000		

表 4.7.3 列點數綜覽—固定製造業、礦業及土石採取業

Overview Row Points^a

地區	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
北部	.250	.767	-.034	.032	.682	.008	1.000	.000	1.000
中部	.250	-.428	-.251	.010	.212	.458	.948	.052	1.000
南部	.250	-.299	.270	.005	.104	.532	.885	.115	1.000
東部金馬	.250	-.040	.014	.000	.002	.001	.935	.019	.954
Active Total	1.000			.048	1.000	1.000			

a. Symmetrical normalization

表 4.7.4 行點數綜覽—固定製造業、礦業及土石採取業

Overview Column Points

營收	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
低營收	.200	-.207	.039	.002	.040	.009	.994	.006	1.000
中低營收	.200	-.321	.100	.005	.096	.059	.984	.015	1.000
中營收	.200	-.554	-.037	.013	.285	.008	.999	.001	1.000
中高營收	.200	.406	-.328	.008	.153	.627	.906	.094	1.000
高營收	.200	.677	.226	.020	.426	.297	.983	.017	1.000
Active Tot	1.000			.048	1.000	1.000			

a. Symmetrical normalization

表 4.7.5 相關測度值—固定製造業、礦業及土石採取業

地區別	營業收入別	相關測度值
北部地區	高	0.1111
北部地區	中高	0.0671
中部地區	中	0.0512
南部地區	中低	0.0214
東部地區及金馬地區	低	0.0018

Row and Column Points

Symmetrical Normalization

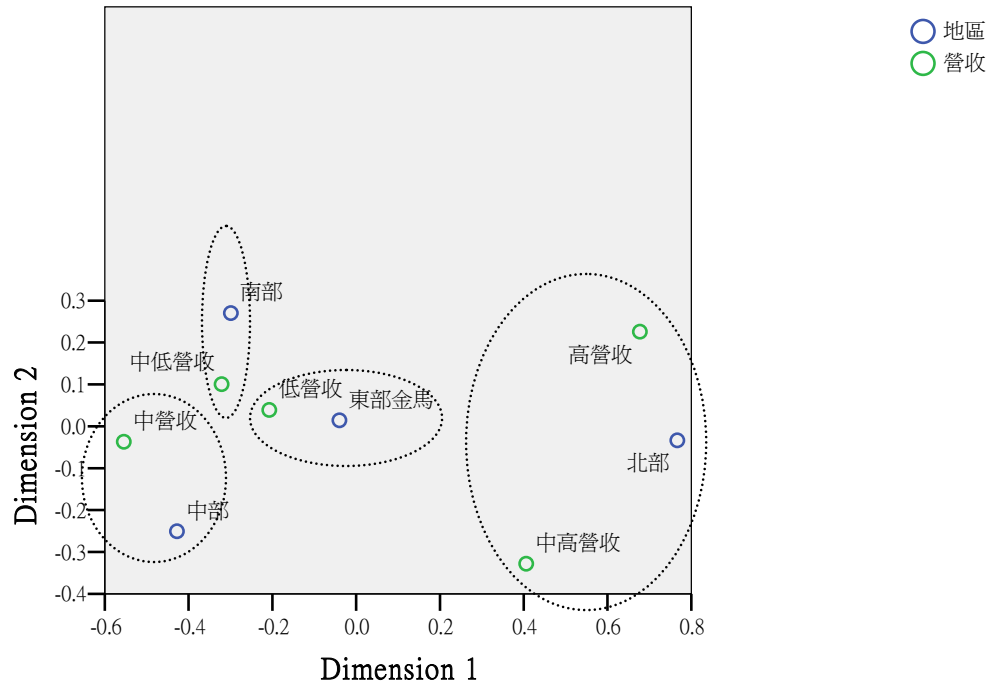


圖 4.7.1 地區與營業收入分布圖—固定製造業、礦業及土石採取業

當行業固定為製造業、礦業及土石採取業時，從圖 4.7.1 各地區與營業收入分層的分布點之距離遠近，可了解高、中高營業收入相關較高的是北部企業，中營業收入相關較高的是中部企業，中低營業收入相關較高的是南部企業，低營業收入相關較高的是東部及金馬地區企業。

為簡化分析，行業固定某一類別，惟研究地區與營業收入類別間之相關強度須考量「行業與地區或行業與營業收入可能存在交互作用關係」，故以【行業×地區】與【行業×營業收入】為新變項進行另一對應分析。當行業分別與地區或營業收入有交互作用時，原 3 維列聯表重新組合成以【行業×地區】與【行業×營業收入】為新變項之(24×30)之 2 維列聯表。

利用表 4.7.6 的 inertia 與表 4.7.7、表 4.7.8 的 score in dimension 求得某兩個類別之間的相關強度：當行業分別與地區或營業收入有交互作用時：

$$\begin{aligned} & \text{【行業} \times \text{地區】與【行業} \times \text{營業收入】相關測度值} = \\ & \sqrt{\dim_1 \text{Inertia}} \times (\text{interaction}_1 \text{ score in dim}_1) \times (\text{interaction}_2 \text{ score in dim}_1) + \\ & \sqrt{\dim_2 \text{Inertia}} \times (\text{interaction}_1 \text{ score in dim}_2) \times (\text{interaction}_2 \text{ score in dim}_2) \end{aligned}$$

藉相同公式，可計算相關測度值如表 4.7.9。

為使對應分析圖中各數據點之標示不致於重疊太多，地區以【N、M、S、E】，營業收入以【低、中低、中、中高、高】，行業別以【製、營、用、商、金、服】簡略表示。圖 4.7.2 顯示於行業類別變動背景（background）下，地區與營業收入類別間之相關強度。表 4.7.9 僅顯示製造業、礦業及土石採取業部分；與表 4.7.5（行業類別固定於製造業、礦業及土石採取業時）比較，相關類別組合及強度順序略有變動，其差異來自比較背景基礎不同。

由表 4.7.7 及表 4.7.8 得知，製造業及礦業及土石採取業、商業及服務業的 score in dimension，相對較大，由圖 4.7.2 觀察，上述 3 個行業，存在明顯差異，且於相同背景（background）下，地區與營業收入類別間之相關強度較大，其中尤以北部地區（N）最突出，其與不同營業收入類別間，分別於對應分析圖 4.7.2 之各項點顯示其相關強度。

表 4.7.6 摘要表—地區、行業及營業收入

Summary						
Dimension	Singular Value	Inertia	Proportion of Inertia		Confidence Singular Value	
			Accounted for	Cumulative	Standard Deviation	Correlation
1	6.012	36.141	.659	.659	.129	.373
2	3.964	15.716	.287	.946	.068	
3	1.060	1.125	.021	.966		
4	.747	.559	.010	.977		
5	.682	.466	.008	.985		
6	.604	.365	.007	.992		
7	.504	.254	.005	.996		
8	.376	.141	.003	.999		
9	.117	.014	.000	.999		
10	.113	.013	.000	.999		
11	.110	.012	.000	1.000		
12	.095	.009	.000	1.000		
13	.078	.006	.000	1.000		
14	.073	.005	.000	1.000		
15	.033	.001	.000	1.000		
16	.019	.000	.000	1.000		
17	.011	.000	.000	1.000		
18	.010	.000	.000	1.000		
19	.007	.000	.000	1.000		
20	.006	.000	.000	1.000		
21	.005	.000	.000	1.000		
22	.003	.000	.000	1.000		
23	.002	.000	.000	1.000		
Total		54.827	1.000	1.000		

表 4.7.7 列點數綜覽—地區、行業及營業收入

Overview Row Points ^a

行業x地區	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
製N	.019	-7.849	5.892	9.805	.190	.162	.699	.260	.959
製M	.019	-5.815	4.316	5.357	.104	.087	.703	.255	.958
製S	.019	-4.748	3.474	3.515	.069	.056	.714	.252	.966
製E	.019	-.167	.121	.005	.000	.000	.564	.197	.761
營N	.019	-.206	-.344	.336	.000	.001	.014	.026	.040
營M	.019	-.101	-.169	.082	.000	.000	.014	.025	.039
營S	.019	-.144	-.240	.178	.000	.000	.013	.024	.037
營E	.019	-.019	-.032	.003	.000	.000	.014	.025	.039
用N	.019	-.166	-.274	.224	.000	.000	.014	.025	.038
用M	.019	-.049	-.081	.022	.000	.000	.012	.022	.034
用S	.019	-.079	-.130	.057	.000	.000	.012	.022	.034
用E	.019	-.008	-.013	.001	.000	.000	.011	.020	.030
商N	.019	-2.247	-9.598	7.682	.016	.430	.073	.880	.954
商M	.019	-.852	-3.567	1.107	.002	.059	.073	.844	.917
商S	.019	-1.240	-5.205	2.388	.005	.127	.072	.833	.904
商E	.019	-.122	-.499	.027	.000	.001	.061	.673	.733
金N	.019	-.202	-.334	.332	.000	.001	.014	.025	.038
金M	.019	-.049	-.082	.020	.000	.000	.014	.024	.038
金S	.019	-.063	-.104	.033	.000	.000	.013	.024	.038
金E	.019	-.006	-.009	.000	.000	.000	.011	.020	.031
服N	.019	10.913	3.049	14.079	.367	.043	.942	.048	.990
服M	.019	5.879	1.702	4.144	.106	.014	.928	.051	.980
服S	.019	6.716	1.949	5.379	.139	.018	.934	.052	.986
服E	.019	.624	.179	.048	.001	.000	.904	.049	.953
Active Total	1.000			54.827	1.000	1.000			

a. Symmetrical normalization

表 4.7.8 行點數綜覽—地區、行業及營業收入

Overview Column Points ^a

行業x營收	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
製低	.018	-.905	.684	.148	.002	.002	.606	.228	.834
製中低	.018	-1.322	.995	.321	.005	.005	.594	.222	.816
製中	.018	-6.803	5.248	7.279	.140	.126	.695	.273	.968
製中高	.018	-7.821	6.157	9.587	.185	.174	.697	.285	.983
製高	.018	-3.106	2.493	1.759	.029	.029	.600	.255	.854
營低	.018	-.040	-.052	.013	.000	.000	.014	.015	.029
營中低	.018	-.063	-.081	.031	.000	.000	.014	.015	.029
營中	.018	-.199	-.260	.330	.000	.000	.013	.015	.028
營中高	.018	-.144	-.190	.191	.000	.000	.012	.014	.026
營高	.018	-.053	-.070	.033	.000	.000	.009	.011	.020
用低	.018	-.041	-.052	.016	.000	.000	.012	.012	.024
用中低	.018	-.049	-.062	.025	.000	.000	.010	.011	.021
用中	.018	-.130	-.165	.167	.000	.000	.011	.012	.023
用中高	.018	-.070	-.090	.060	.000	.000	.009	.010	.019
用高	.018	-.035	-.046	.018	.000	.000	.007	.008	.016
商低	.018	-.696	-2.825	.727	.001	.037	.073	.792	.865
商中低	.018	-.563	-2.324	.457	.001	.025	.076	.853	.929
商中	.018	-1.638	-6.855	3.860	.008	.215	.076	.877	.953
商中高	.018	-1.709	-7.455	4.609	.009	.255	.069	.869	.939
商高	.018	-.650	-2.918	.798	.001	.039	.058	.769	.827
金低	.018	-.045	-.057	.019	.000	.000	.011	.012	.024
金中低	.018	-.031	-.040	.010	.000	.000	.011	.012	.023
金中	.018	-.094	-.121	.094	.000	.000	.010	.011	.021
金中高	.018	-.108	-.141	.131	.000	.000	.010	.011	.021
金高	.018	-.074	-.099	.091	.000	.000	.007	.008	.014
服低	.018	6.725	2.105	5.319	.137	.020	.929	.060	.990
服中低	.018	8.782	2.759	9.048	.233	.035	.932	.061	.992
服中	.018	8.895	2.813	9.297	.239	.036	.930	.061	.992
服中高	.018	1.525	.498	.353	.007	.001	.721	.051	.771
服高	.018	.462	.152	.037	.001	.000	.631	.045	.676
Active Total	1.000			54.827	1.000	1.000			

a. Symmetrical normalization

表 4.7.9 相關測度值—行業為製造業、礦業及土石採取業

行業×地區別	行業×營業收入別	相關測度值
行業×北部地區	行業×高營業收入	204.79
行業×北部地區	行業×中高營業收入	512.85
行業×中部地區	行業×中營業收入	327.61
行業×南部地區	行業×中低營業收入	51.43
行業×東部地區及金馬地區	行業×低營業收入	1.24

Row and Column Points

Symmetrical Normalization

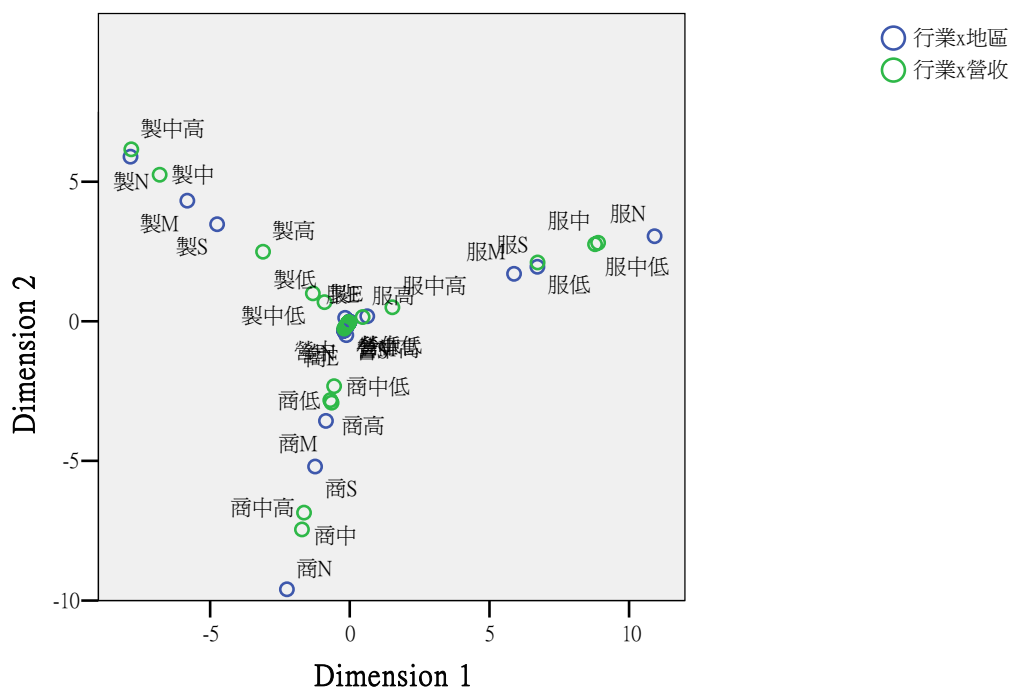


圖 4.7.2 地區、行業及營業收入分布圖

第八節 結果與討論

茲依據本章前揭各節內容，歸納彙整分析結果及相關問題如次：

一、 樣本代表性分析：

- (一) 母體及樣本資料營運特徵值之行業分布：由表 4.1.1、表 4.2.1 及表 4.3.1 觀察家數比率最高之製造業、批發及零售業與住宿及餐飲業三大產業，母體以批發及零售業占 46.00%居首，高於製造業之 13.90%與住宿及餐飲業之 8.33%；惟該業樣本家數占 19.34%，而住宿及餐飲業更萎縮至 1.30%，製造業則占 28.93%，另從業員工人數、資產淨額及營業收入等重要營運特徵值之母體及樣本行業結構亦有所差異。
- (二) 各大類行業樣本抽出率：95 年普查係考量層內變異，採紐曼法配置樣本，樣本總抽出率為 6.85%(含全查層)，如排除分工調查及家數較少之行業部分，以製造業抽出率 14.27%最高，住宿及餐飲業抽出率 1.07%最低，各大類行業之抽出率差異頗大。
- (三) 行業改變及空層情形：觀察 95 年普查行業改變情形，原抽樣廠商之細類行業，與回表行業不同達 35.62%；復以未回表情形仍存，故除全查樣本外，空層數及空層率分別達 329 層及 8.94%，其中以製造業與批發及零售業空層數分別為 150 層與 53 層最多，至空層率以藝術、娛樂及休閒服務業(含 11 個細類行業，樣本 743 家，抽出率 4.15%) 21.05%最高。
- (四) 歷次普查母體資料結構變化：以「營業收入」為準則變數，藉 CHAID 分析普查母體各項變數組合的營業收入，發現「從業員工人數」、「總支出」及「從業員工薪資」，20 年來均屬重要因素；至「細類行業」則自 85 年起，重要性趨弱；「資產淨額」與「縣市鄉鎮」重要性較低，致分別於 95 年及 90 年退出自變數行列。

二、**樣本精確度**：因母體與樣本資料之重要特徵值多呈明顯右偏，爰將營業收入、從業員工人數及資產淨額 3 項特徵值，取 0.134912 次方轉換後截尾，以細類業別平均數為基礎，比較樣本與母體之差異，其中資產淨額部分，逾 5 成之細類行業樣本均數超過母體均數之 3 倍；至營業收入及從業員工人數，其樣本均數超過母體均數 3 倍之細類行業亦分別達 47.20% 及 23.86%。各細類行業中，以電視傳播業(6021)之樣本精確度最佳。

三、**樣本特性分析**：將 95 普查抽樣作業應用之相關重要特徵值之關聯性分析結果列述如次：

(一)**列聯表分析**：由卡方獨立性檢定結果及 Cramer's V 統計量得知，「大類行業」與「地區別」之樣本家數分布，存在顯著差異。觀察各大類行業樣本家數之地區分布情形，除「礦業及土石採取業」樣本家數分布以中部地區 36.7% 居首外，餘大類行業分布皆側重北部地區，樣本家數結構比率最低者(用水供應及污染整治業)亦逾 4 成(40.2%)；至東部地區及金馬地區，樣本家數分布顯著低於其他地區，合計僅占 2.2%。

(二)**關聯分析**：經由 CHAID 分層後，觀察「印刷業」與「服裝及其配件批發業」二細類行業特徵值之關聯性得知，各連續變數及類別變數中，各層「營業收入」皆與「總支出」(相關係數皆在 0.97 以上)及「縣市別」(E^2 皆在 0.2266 以上)之相關程度最高；至各層類別變數以「縣市別」預測時，以「組織別」可消滅的錯誤較多(除「服裝及其配件批發業」第 8 層外，各層 Lambda 值皆在 0.125 以上)。

(三)**對數線性模式分析**：為觀察影響樣本家數分布總重要之變項，爰以 95 年抽樣作業重要變數—「營業收入」、「地區別」及「行業別」為模式之 3 項效應，進行對數線性模式分析，由飽和及非飽和模型各效應檢定結果得知，單一效應及效應間之交互作用，檢定結果皆屬顯著；惟因各非飽和模型之適合度檢定亦皆

呈顯著，可知其效應之間存在某種程度之關係，但此關係不夠強烈到足以構成一個完整模型的程度。為求得各效應解釋樣本家數分布之強度，爰計算各效應之「最大資訊估計量」，並得知「行業別」為解釋比率最大之效應(40%)，其次為「營業收入」(39%)及「地區別」(21%)。

(四) **對應分析**：由對數線性模式分析可得知各效應對於樣本家數分布之解釋強度，復進一步藉對應分析觀察 3 項效應間之關聯性。當行業固定為「製造業」與「礦業及土石採取業」時，北部地區與高營業收入及中高營業收入、中部地區與中營業收入、南部地區與中低營業收入、東部地區及金馬地區與低營業收入之相關測度值皆為正，其中以北部企業與高營業收入之關係最強(相關測度值為 0.1111)，而以東部企業與低營業收入之關係最弱(相關測度值為 0.0018)。如綜合觀察三項效應之關聯性(即同時考慮所有行業類別)，則結果大致相同，僅效應組合及相關之強弱順序略有變動。

(五) **CART 分析**：為觀察各財務報表細項資料或營運特徵值與營業收入之關聯性，爰以製造業為例，自問項挑選 26 個變項為自變數，進行樣本資料之 CART 分析，由結果得知，對「營業收入」言，以「中間消費」及「從業員工薪資」為最重要的變數。而製造業之中間消費=原材物料及燃料耗用總值+水電瓦斯費+託外加工費+其他營業費用，其中以「原材物料及燃料耗用總值」占逾 8 成(82.19%)最高，顯示該變項係影響營業收入高低之重要變項。

四、 **研究限制**：辦理 95 年普查名冊整理作業時，須執行新舊行業轉換(由中華民國行業標準分類第 7 次修訂轉換為第 8 次修訂)，致新興行業如信託服務業(6632)、企業總管理機構(7010)、非金融性無形資產租賃業(7740)、會議及展覽服務業(8202)及複合支援服務業(8110)等 5 個細類行業因無母體資料，無法抽樣及推估；資料儲存媒體複製業(1620，母體企業計 11 家)與資料儲存媒體製

造業(2740，母體企業 116 家)未明顯區隔，致資料儲存媒體複製業(1620)無法掌握母體資料而未抽取樣本。除信託服務業(6632)、企業總管理機構(7010)無企業資料外，餘 4 個細類行業須依前揭空層處理方式，藉相近細類行業資料作為推估基礎，經事後分層後，進行推計作業。

第五章 結論與建議

第一節 結論

茲就本研究之實證分析結果，探討普查抽樣作業之效能，並引申全文結論如次：

一、**普查分層方法之選取**：近年來產業變遷迅速，企業經營模式已跳脫傳統思維，其獲利能力未必與投入存在絕對之比例關係；而綜合樣本特性分析結果，影響普查資料分布及其特性之特徵值實愈趨多元，運用單一變數衡量企業經營效能，將無法有效區分產業之特性；另產業經營特徵值分布偏斜之特質，亦增加相關研究分析之困擾。本次普查分層方法引入多變數分層概念，且因考量特徵變數間之互動性，降低層內變異之效果甚佳，使分層結果愈趨穩健，樣本規模亦隨之降低。95年普查以多變數分層，符合普查資料之特性，有助於掌握母體資料結構，並抽出具代表性廠商，同時可兼顧多變數之互動關係，以達隨機抽樣的效果。

二、**普查之樣本精確度**：

(一)**精確度分析結果**：由各細類行業樣本之特徵值均數，相較於母體，皆發生偏高情形，其中又以資產淨額之差異最大。而各細類行業中差異情形較大者，皆屬樣本抽出率較小行業(如零售業部分細類行業)；而抽出率愈高者(如製造業部分細類行業，及分工調查之相關行業等)，精確度愈佳。此一結果雖淺顯且可預期，惟為提高樣本之精確性，確應審酌樣本規模之妥適性。

(二)**樣本與母體結構差異**：綜合前章分析結果，母體及樣本結構確有不一致之情形。除全查層、特殊行業及分工機關調查廠商另作處理外，餘採分細類行業分層抽樣，各層樣本數係採紐曼法配置。惟多變數分層內涵已考量特徵變數間之互動關係，及層內變異小、層間變異大之特質，致各細類行業各層推估確度不一致。

(三) 行業分類及空層情形：由分析結果得知，樣本抽出率相對較高之細類行業，仍發生樣本平均特徵值明顯高於母體情形，部分原因係該細類行業發生回表行業與原抽樣行業不一致，以及空層率偏高所致。因企業經營模式愈趨複雜，行業分類若同質性過高或過細，或抽樣母體未能即時更新而與現況不符，將無法反映企業之多角化經營模式，致調查行業與回表行業不同，甚至造成整層樣本皆未回表之空層情形；而因層內完全無樣本，必須藉他層資訊作為推估基礎，對於推估確度實有影響。

三、 普查分層之變數選取：

由製造業特性得知，「中間消費」及「從業員工薪資」為影響營業收入重要之變項，其中尤以「中間消費」影響最大。中間消費係廠商從事生產活動時所需之各項成本投入，包括相關部門所提供之產品及非要素勞務(含服務)，為普查多個問項之資料加總而得。抽樣之分層特徵值雖未包括中間消費，惟其與資產淨額、總支出及總收入皆存在顯著之高相關性；另從業員工薪資與人數亦存在高度關聯。抽樣作業以資產淨額、從業員工及縣市鄉鎮為分層特徵變項，進而解釋企業之營業收入，確可臻分層之預期效益。

以印刷業、服裝及其配件批發業為例，企業之營業收入與總支出相關程度最高，與地區亦有所關聯，組織型態亦可由地區別明顯反映，顯示此 2 細類行業因經營模式及型態較為傳統，附加價值偏低，且侷限於固定區域；另地區、行業及營業收入之交互作用顯著，顯示大部分產業皆側重於北部地區及製造業，產業集中趨勢明顯，區域性產業發展未能平衡之情形仍存。而此 3 變項皆係解釋廠商家數分布之重要效應，抽樣作業以細類行業為副母體，並以營業收入為準則變數，地區別為分層方法之自變數，確與本分析結果相符。另由樣本資料之對應分析結果，適足印證普查結果中，廠商營業收入及地區分布情形，亦再次反映普查分層作業相關變數選取之妥適性。

第二節 建議

依前節所述，分層方法及相關變數之選取，雖經實證分析，具備足夠之妥適性；惟調查行業與回表行業不同及空層情形，致改進空間仍存。為降低推計偏誤，提升資料確度，爰提出改進方針如次：

- 一、**衡酌成本及確度，制訂抽樣規模**：降低廠商填報負擔，係近年企業面調查辦理之基本原則，惟抽樣規模之制訂牽涉層面極廣，雖於普查前請縣市政府主計室人員確認抽樣廠商是否持續營業或符合普查對象，並以候補樣本或加抽樣本等方式處理，惟較上次普查抽樣規模大幅下降近 10 萬家廠商，由分析結果顯示，實存在樣本推估確度不足之現象。爰建議以精確度最佳化為前提，衡酌經費及人力情形，制訂妥適之抽樣規模。以紙容器製造業為例，本次普查母體共 2,549 家，樣本數共 193 家，若採比例配置各層樣本數，其抽樣誤差為 6.86%，若抽樣誤差降低至 5% 以下，樣本數至少須增加 49.74%(96 家)。
- 二、**行業分類及空層情形**：普查樣本資料之行業及空層情形，對於推計確度之影響甚大，行業歸類原則及方式，爰須通盤考量普查資料與企業現況，避免同質性過高或過細；另亦應研擬合併副母體之可行性，俾降低分類過細或母體家數過少而致空層或回表行業移轉之情形發生。至訪查執行方面，建議運用基層統計調查網人員，由原電話確認改採實地判定作業方式，俾確實先期掌握廠商持續營運情形，降低未回表情形發生；並應於訪查期間，隨時掌握各副母體各層回表狀況，如發現偏低情形，即時通報縣市人員追究原因並予補救，俾維持樣本結構，提升樣本之精確度。
- 三、**採比例配置法配置各層樣本**：多變數分層內涵，實已考量特徵變數間之互動關係，與層間變異大、層內變異小之特質，惟本次普查採紐曼配置法決定各層樣本數，因層內變異重複考量，致影響抽樣效能。建議改採比例配置法作業，除符合多變數分層理論基礎外，整體樣本結構亦較能充分反映母體，並彌補原紐曼配置降

低小規模層樣本數，致精確度不足之缺憾。

四、**依副母體特性選取分層變數**：本次普查各副母體(細類行業)係依不同分層方法，選取固定之分層變數(母體家數 1 千家以上者以營業收入為準則變數，從業員工人數、資產淨額、縣市鄉鎮為特徵變數；家數未滿 1 千家則皆以營業收入及從業員工人數為分層變數)，並未考量各業別間之差異性，致部分特殊行業特徵值層內變異之降低效果較不顯著(如民間融資業家數未滿 1 千家，未考慮資產分布，致層內之資產變異仍大)。建議針對個別副母體之各項特徵值，運用妥適方法，分別選取分層變數，俾落實降低層內變異之原則。

五、**即時性母體檔之建置及未來展望**：

(一)**即時性母體檔之建置**：廠商即時營運資訊之取得，將有助於抽樣效能及資料確度之提升，預期除可降低樣本回表行業與原抽樣行業不一致之情形外，抽樣作業及分層效果亦將因相關變數之即時更新，而獲致改善。本處將自 99 年度起，以最近一次普查資料為基礎，定期蒐集各機關公務登記及抽樣調查資訊，常川更新普查資料，建置即時性母體檔供各界應用，俾建立工商普查及各機關抽樣調查雙向回饋之良性機制。

(二)**未來展望**：為期永續發展我國工商業統計資訊，就長期規劃言，建議參酌各國整合性經濟調查模式，建立機關協商、分工及作業統合機制，常川性蒐集各行業各類工商業經營資訊，取代目前 5 年辦理經濟普查之方式；並運用資料採礦技術，建置定期更新之母體資料庫，提供便捷資訊擷取平臺及各機關辦理抽樣調查之母體即時資訊，俾臻降低成本、節省人力及提升資料確度之多重目的。

參考文獻（中文部分）

1. 呂金河編譯(2005)，多變量分析，滄海書局。
2. 行政院主計處(2006)，95年工商及服務業普查報告，第36卷普查規制與作業方法。
3. 徐村和(2000)，「信用卡模糊市場區隔之研究」，輔仁管理評論，第七卷第二期，1-20。
4. 陳丁群(2006)，「以治病基因集為先驗資訊的基因選取方法之研究」，國立成功大學資訊管理研究所碩士論文。
5. 陳正昌、程炳林、陳新豐、劉子鍵(2007)，多變量分析方法—統計軟體應用，五南圖書出版股份有限公司。
6. 黃登源(1998)，應用迴歸分析，華泰文化事業公司。
7. 黃登源(2006)，應用多變量分析講義。
8. 黃登源，李仁棻(2009)，統計資料分析講義。
9. 黃登源，李仁棻(2009)，應用統計講義。
10. 黃麗蒨(2000)，「臺灣地區地下水品質之統計研究」，國立中央大學統計研究所碩士論文。
11. 樓邦儒、賈立人(2000)，「柯本氣候分類與模糊氣候分類之比較—以台灣為例」，國立臺北師範學院學報，第十三期(八十九年六月)293~310。

參考文獻（英文部分）

1. A.R. Bagnall(1985) , ”Integration of Economic Statistics : The Conceptual and Operational Framework” , Australian Bureau of Statistics , Australia.
2. Agresti, A.(2002) , Categorical Data Analysis , 2nd ed. New York : Wiley.
3. Andrea Guizzardi(2004) , ”A New Approach in Business Travel Survey:Multivariate Techniques for Strata Design” , Department of Statistical Science , University of Bologna, Italy.
4. Australian Bureau of Statistics(2003) , Strategies for Synthetic Estimation in ABS Business Surveys , Australia.
5. Australian Bureau of Statistics(2005) , Estimation Methodologies Using Taxation Data for ABS Business Surveys , Australia.
6. Daniela Cocchi 、 Enrico Fabrizi(2002) , ”Regression Trees based Stratification : an Application to the Analysis of the Italian Post Enumeration Survey” , Meri Raggi and Carlo Trivisano Dipartimento di Scienze Statistiche , Università di Bologna , Italia.
7. Deng-Yuan HUANG 、 Ren-Fen LEE(2003) , ”On some data orient robust estimate procedure for means” , Journal of Applied Statitics , Vol.30 , No.6 , July 2003 , pp.625-634.
8. Deng-Yuan HUANG 、 Ren-Fen LEE 、 S.Pachapakesan(2006) , ”On some variable selection procedures for regression models” , Journal of Statistical Planning and Inference , Vol.136 , Issue7 , July 2006 , pp.2020-2034.
9. Hanna Elisson and Eva Elvers(2001) , “Cut-Off Sampling and Estimation” , Statistics Sweden.

10. International Benchmark National Accounts(2007) , Statistics Netherlands.
11. Jay Fowdar 、Zuhair Bandar 、Keeley Crockett(2002) , ”Inducing Fuzzy Decision Trees in Non-Deterministic Domains using CHAID” , Department of Computing and Mathematics , Manchester Metropolitan United Kingdom.
12. M. Ballin e G. Barcaroli(2008) , ”Stratification of Sampling Frames in a Multivariate and Multidomain Sample Design” , Istat , Italy.
13. Macro Bee 、Roberto Benedetti 、Giuseppe Espa(2007) , ”A Framework for Cut-Off Sampling in Business Survey Design” , Department of Economics , University of Trento , Italy.
14. Methodology for Producing Synthetic Microdata for Income in Non-Survey Years(2004) , Australian Bureau of Statistics , Australia.
15. Pascal Rivière (1999) , ”Coodinating Samples Using the Microstrata Methodology” , INSEE , France.
16. Peter G.M. Van Der Heijden, Jan De Leeuw (1985) , ”Correspondence Analysis Used Complementary to Log-linear Analysis” , Psychometrika Netherlands , VOL. 50 , NO.4 , DEC. 1985 , pp.429~427.
17. Statistics Canada(2006) , The Integrated Approach to Economic Surveys in Canada.
18. Stephen Penneck(2005) , ”Developing integrated economic statistics systems - learning from the past” , paper for UN workshop on Economic Censuses , ONS , United Kingdom.
19. United Nations Statistics Division (2007) , The Results of the UNSD Survey of Country Practices in Economic Census.
20. William E. Winkler(1998) , Strata Boundary Determination , U.S. Census Bureau.

附錄 1 連續營運特徵值間的相關係數—印刷業

	變數名稱	變數名稱	相關係數	P-Value
第 1 層	從業員工人數	樓地板面積	0.87857	0.0092
	從業員工人數	總支出	0.84226	0.0174
	從業員工薪資	勞動報酬	0.92840	<.0001
	從業員工薪資	總支出	0.91924	0.0034
第 2 層	土地面積	樓地板面積	0.98207	<.0001
	從業員工人數	勞動報酬	0.93939	0.0005
	從業員工人數	從業員工薪資	0.90959	0.0017
	勞動報酬	存貨及存料資產	0.76706	0.0263
第 3 層	土地面積	樓地板面積	0.97170	<.0001
	從業員工人數	房屋及建築淨額資產	0.99142	<.0001
	從業員工人數	市場行銷金額	0.84533	0.0005
	從業員工薪資	勞動報酬	0.98710	<.0001
第 4 層	從業員工薪資	總支出	0.86711	0.0003
	勞動報酬	總支出	0.92604	<.0001
	從業員工薪資	勞動報酬	0.97181	<.0001
	從業員工薪資	總支出	0.90488	0.0020
第 5 層	勞動報酬	總支出	0.96550	<.0001
	樓地板面積	房屋及建築淨額資產	0.93395	0.0007
	市場行銷金額	存貨及存料資產	0.88345	0.0036
	勞動報酬	從業員工薪資	0.99770	<.0001
第 5 層	從業員工人數	勞動報酬	0.80909	0.0003
	從業員工人數	從業員工薪資	0.78259	0.0006
	從業員工人數	總支出	0.78179	0.0006
	從業員工人數	土地面積	0.58414	0.0222
	從業員工薪資	土地面積	0.69281	0.0042
	從業員工薪資	總支出	0.64486	0.0094
	勞動報酬	土地面積	0.70602	0.0033
	勞動報酬	總支出	0.65782	0.0077
	樓地板面積	土地面積	0.62316	0.0131
	樓地板面積	土地資產	0.60757	0.0163
	樓地板面積	存貨及存料資產	0.58408	0.0222
	樓地板面積	房屋及建築淨額資產	0.51925	0.0473
	第 6 層	從業員工人數	從業員工薪資	0.86057
從業員工人數		勞動報酬	0.82900	0.0002
從業員工人數		存貨及存料資產	0.76809	0.0013
從業員工人數		資產淨額	0.55105	0.0411
從業員工薪資		勞動報酬	0.99166	<.0001
從業員工薪資		存貨及存料資產	0.74608	0.0022
從業員工薪資		資產淨額	0.76453	0.0014
從業員工薪資		總支出	0.80720	0.0005
勞動報酬		總支出	0.83289	0.0002
勞動報酬		存貨及存料資產	0.70960	0.0045
勞動報酬		資產淨額	0.77581	0.0011
總支出		土地資產	0.55672	0.0387
總支出		資產淨額	0.69795	0.0055
土地面積	樓地板面積	0.87933	<.0001	
第 7 層	從業員工人數	從業員工薪資	0.70688	0.0005
	從業員工人數	勞動報酬	0.68012	0.001
	從業員工人數	市場行銷金額	0.64153	0.0023
	從業員工薪資	勞動報酬	0.99291	<.0001
	從業員工薪資	員工訓練金額	0.55597	0.0109
	從業員工薪資	資產淨額	0.53507	0.0151
	從業員工薪資	總支出	0.51410	0.0204
	從業員工薪資	土地資產	0.51363	0.0205
	勞動報酬	員工訓練金額	0.54794	0.0124
	勞動報酬	總支出	0.48834	0.0289
	勞動報酬	土地資產	0.31491	0.0230
	勞動報酬	資產淨額	0.49798	0.0255
	總支出	存貨及存料資產	0.74879	0.0001
土地面積	樓地板面積	0.55924	0.0104	
第 8 層	從業員工人數	從業員工薪資	0.81027	<.0001
	從業員工人數	勞動報酬	0.83071	<.0001
	從業員工人數	總支出	0.71493	<.0001

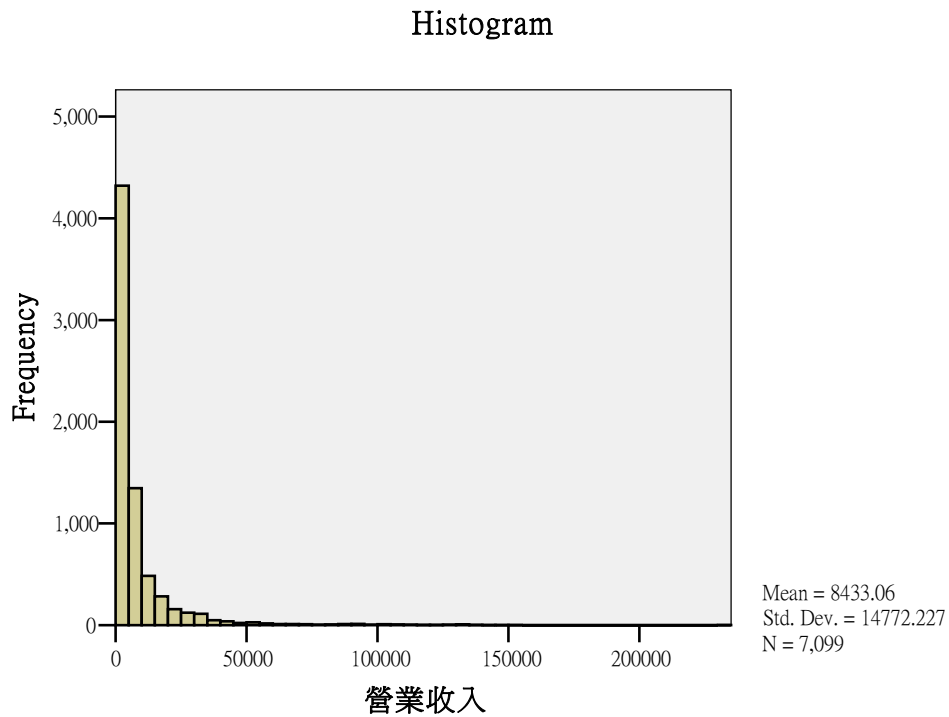
	變數名稱	變數名稱	相關係數	P-Value
	從業員工人數	存貨及存料資產	0.53419	0.0014
	從業員工薪資	勞動報酬	0.99378	<.0001
	從業員工薪資	總支出	0.73317	<.0001
	從業員工薪資	存貨及存料資產	0.57606	0.0005
	勞動報酬	總支出	0.72246	<.0001
	勞動報酬	存貨及存料資產	0.55641	0.0008
	勞動報酬	資產淨額	0.40430	0.0196
	總支出	存貨及存料資產	0.68668	<.0001
	土地面積	樓地板面積	0.93151	<.0001
第 9 層	從業員工人數	從業員工薪資	0.89424	<.0001
	從業員工人數	勞動報酬	0.88654	<.0001
	從業員工人數	總支出	0.74929	<.0001
	從業員工人數	資產淨額	0.58154	<.0001
	從業員工人數	樓地板面積	0.38468	<.0001
	從業員工人數	土地面積	0.34436	<.0001
	從業員工人數	土地資產	0.40960	<.0001
	從業員工人數	房屋及建築淨額資產	0.40488	<.0001
	從業員工人數	運輸設備淨額資產	0.32053	0.0002
	從業員工人數	市場行銷金額	0.27364	0.0016
	從業員工人數	電腦軟體資料庫金額	0.26954	0.0018
	從業員工人數	存貨及存料資產	0.23915	0.0059
	從業員工人數	員工訓練金額	0.23432	0.0071
	從業員工薪資	勞動報酬	0.99385	<.0001
	從業員工薪資	總支出	0.78874	<.0001
	從業員工薪資	資產淨額	0.59345	<.0001
	從業員工薪資	土地面積	0.44831	<.0001
	從業員工薪資	樓地板面積	0.53093	<.0001
	從業員工薪資	房屋及建築淨額資產	0.35053	<.0001
	從業員工薪資	土地資產	0.36089	<.0001
	從業員工薪資	市場行銷金額	0.30082	0.0005
	從業員工薪資	運輸設備淨額資產	0.27213	0.0017
	從業員工薪資	電腦軟體資料庫金額	0.26353	0.0024
	從業員工薪資	員工訓練金額	0.30082	0.0040
	從業員工薪資	研究發展金額	0.26353	0.0131
	從業員工薪資	存貨及存料資產	0.19657	0.0244
	勞動報酬	總支出	0.80596	<.0001
	勞動報酬	資產淨額	0.58869	<.0001
	勞動報酬	房屋及建築淨額資產	0.33843	<.0001
	勞動報酬	土地資產	0.34923	<.0001
	勞動報酬	土地面積	0.41979	<.0001
	勞動報酬	樓地板面積	0.50689	<.0001
	勞動報酬	運輸設備淨額資產	0.29201	0.0007
	勞動報酬	市場行銷金額	0.22467	0.0013
	勞動報酬	電腦軟體資料庫金額	0.27087	0.0018
	勞動報酬	研究發展金額	0.24210	0.0053
	勞動報酬	員工訓練金額	0.22467	0.0099
	勞動報酬	存貨及存料資產	0.19039	0.0294
	總支出	電腦軟體資料庫金額	0.33982	<.0001
	總支出	樓地板面積	0.35514	<.0001
	總支出	土地資產	0.37853	<.0001
	總支出	資產淨額	0.67412	<.0001
總支出	存貨及存料資產	0.32824	0.0001	
總支出	房屋及建築淨額資產	0.33139	0.0001	
總支出	土地面積	0.31724	0.0002	
總支出	市場行銷金額	0.29142	0.0007	
總支出	運輸設備淨額	0.26115	0.0026	
總支出	研究發展金額	0.23831	0.0061	
總支出	員工訓練金額	0.20162	0.0209	
資產淨額	土地面積	0.31724	0.0001	
資產淨額	樓地板面積	0.35514	<.0001	
土地面積	樓地板面積	0.70767	<.0001	

附錄 2 連續營運特徵值間的相關係數—服裝及其配件批發業

層別	變數名稱	變數名稱	相關係數	P-Value
第 1 層	從業員工人數	總支出	0.76018	0.0473
	從業員工人數	員工訓練金額	0.98622	<.0001
	從業員工人數	電腦軟體資料庫金額	1	<.0001
	從業員工人數	存貨及存料資產	0.91018	0.0044
	從業員工薪資	勞動報酬	0.99306	<.0001
	總支出	電腦軟體資料庫金額	0.76018	0.0473
	資產淨額	樓地板面積	0.87171	0.0106
第 2 層	從業員工人數	從業員工薪資	0.72118	0.0024
	從業員工人數	勞動報酬	0.70249	0.0035
	從業員工人數	總支出	0.56359	0.0287
	從業員工人數	房屋及建築淨額資產	-0.54137	0.0371
	從業員工薪資	勞動報酬	0.98771	<.0001
	勞動報酬	總支出	0.94001	<.0001
	從業員工薪資	總支出	0.91553	<.0001
	從業員工薪資	資產淨額	0.56794	0.0272
	勞動報酬	資產淨額	0.55344	0.0323
	土地面積	樓地板面積	0.84841	<.0001
第 3 層	從業員工人數	市場行銷金額	0.81325	<.0001
	從業員工人數	從業員工薪資	0.75734	<.0001
	從業員工人數	勞動報酬	0.74708	<.0001
	從業員工薪資	勞動報酬	0.99784	<.0001
	從業員工薪資	總支出	0.67549	0.0006
	勞動報酬	總支出	0.67729	0.0005
	總支出	中間消費	0.79162	<.0001
	總支出	資產淨額	0.49604	0.0189
	土地面積	樓地板面積	0.88521	<.0001
第 4 層	從業員工薪資	勞動報酬	0.97851	<.0001
	從業員工薪資	總支出	0.54969	0.0148
	勞動報酬	總支出	0.52288	0.0216
	土地面積	樓地板面積	0.98799	<.0001
第 5 層	勞動報酬	從業員工薪資	0.97863	0.0037
	土地面積	樓地板面積	0.9677	0.0069
第 6 層	從業員工人數	從業員工薪資	0.96353	0.0083
	從業員工人數	勞動報酬	0.96647	0.0073
	從業員工人數	機械及什項設備資產	0.92136	0.0262
	從業員工人數	中間消費	0.91712	0.0283
	從業員工薪資	勞動報酬	0.99947	<.0001
	從業員工薪資	機械及什項設備資產	0.9392	0.0178
	從業員工薪資	中間消費	0.98279	0.0027
	勞動報酬	機械及什項設備資產	0.94283	0.0163

層別	變數名稱	變數名稱	相關係數	P-Value
	勞動報酬	中間消費	0.97721	0.0041
	總支出	存貨及存料資產	0.90408	0.0351
	資產淨額	市場行銷金額	0.95436	0.0116
第 7 層	從業員工人數	從業員工薪資	0.42788	0.0047
	從業員工人數	勞動報酬	0.41563	0.0062
	從業員工人數	資產淨額	0.37383	0.0147
	從業員工人數	樓地板面積	0.3589	0.0196
	從業員工人數	機械及什項設備資產	0.311	0.045
	從業員工薪資	勞動報酬	0.99164	<.0001
	從業員工薪資	中間消費	0.50254	0.0007
	從業員工薪資	總支出	0.60928	<.0001
	從業員工薪資	專業技術購入金額	0.37219	0.0152
	從業員工薪資	資產淨額	0.37003	0.0159
	從業員工薪資	樓地板面積	0.33216	0.0316
	勞動報酬	總支出	0.59276	<.0001
	勞動報酬	資產淨額	0.39514	0.0096
	勞動報酬	專業技術購入金額	0.38227	0.0125
	勞動報酬	樓地板面積	0.32374	0.0365
	總支出	中間消費	0.77382	<.0001
	總支出	專業技術購入金額	0.39694	0.0092
	總支出	資產淨額	0.38359	0.0122
	土地面積	樓地板面積	0.4925	0.0009
	第 8 層	從業員工人數	從業員工薪資	0.86751
從業員工人數		勞動報酬	0.85878	<.0001
從業員工人數		專業技術購入金額	0.46123	0.0001
從業員工人數		專業技術銷售金額	0.33372	0.0066
從業員工薪資		勞動報酬	0.99728	<.0001
從業員工薪資		專業技術購入金額	0.50647	<.0001
從業員工薪資		資產淨額	0.39842	0.001
勞動報酬		機械及什項設備資產	0.38966	0.0013
從業員工薪資		機械及什項設備資產	0.36732	0.0026
從業員工薪資		總支出	0.26196	0.035
勞動報酬		專業技術購入金額	0.48619	<.0001
勞動報酬		資產淨額	0.42234	0.0005
勞動報酬		總支出	0.29823	0.0158
總支出		中間消費	0.73262	<.0001
總支出		資產淨額	0.44933	0.0002
土地面積		樓地板面積	0.57367	<.0001
中間消費		樓地板面積	0.26593	0.0323

附錄5 印刷業(1611)營業收入分布—母體資料

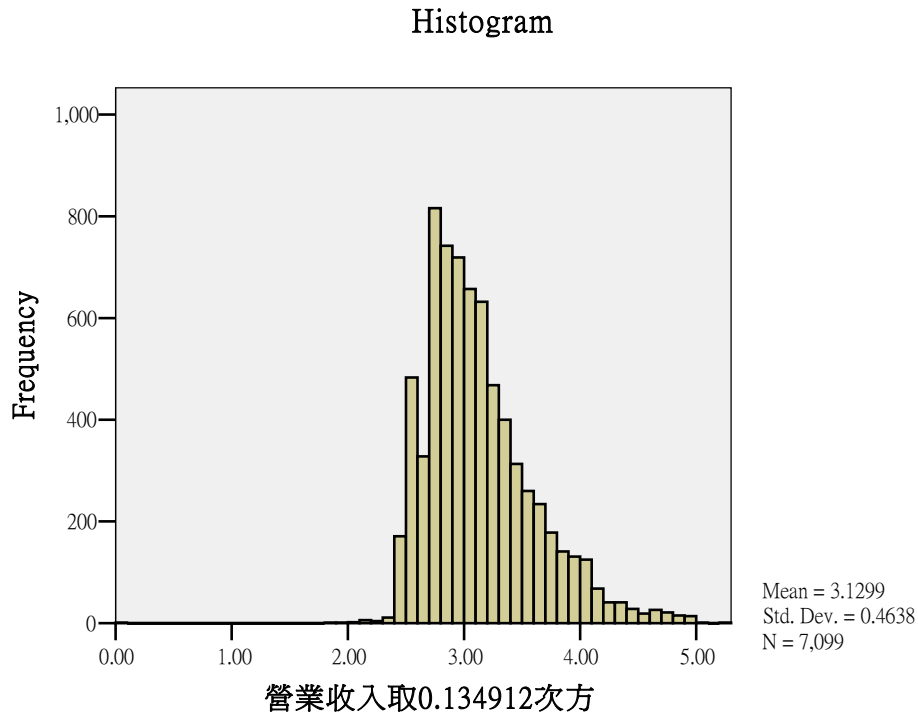


營業收入 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
377.00	0 .	78888899999&
1302.00	1 .	000001111222334455666777778888899999
1240.00	2 .	0000001111222223334445556666777788889
807.00	3 .	000111223344555667778889
595.00	4 .	0011223445566778899
429.00	5 .	00112344567789
303.00	6 .	0123456789
258.00	7 .	0123456789
175.00	8 .	03467&&
181.00	9 .	034589&&
116.00	10 .	&&&&
118.00	11 .	&&&&
90.00	12 .	4&&
75.00	13 .	&&
86.00	14 .	&&&
76.00	15 .	&&
66.00	16 .	&&
5.00	17 .	&
800.00	Extremes	(>=17111)

Stem width: 1000
Each leaf: 33 case(s)
& denotes fractional leaves.

附錄6 印刷業(1611)營業收入分布—母體資料取0.134912次方轉換

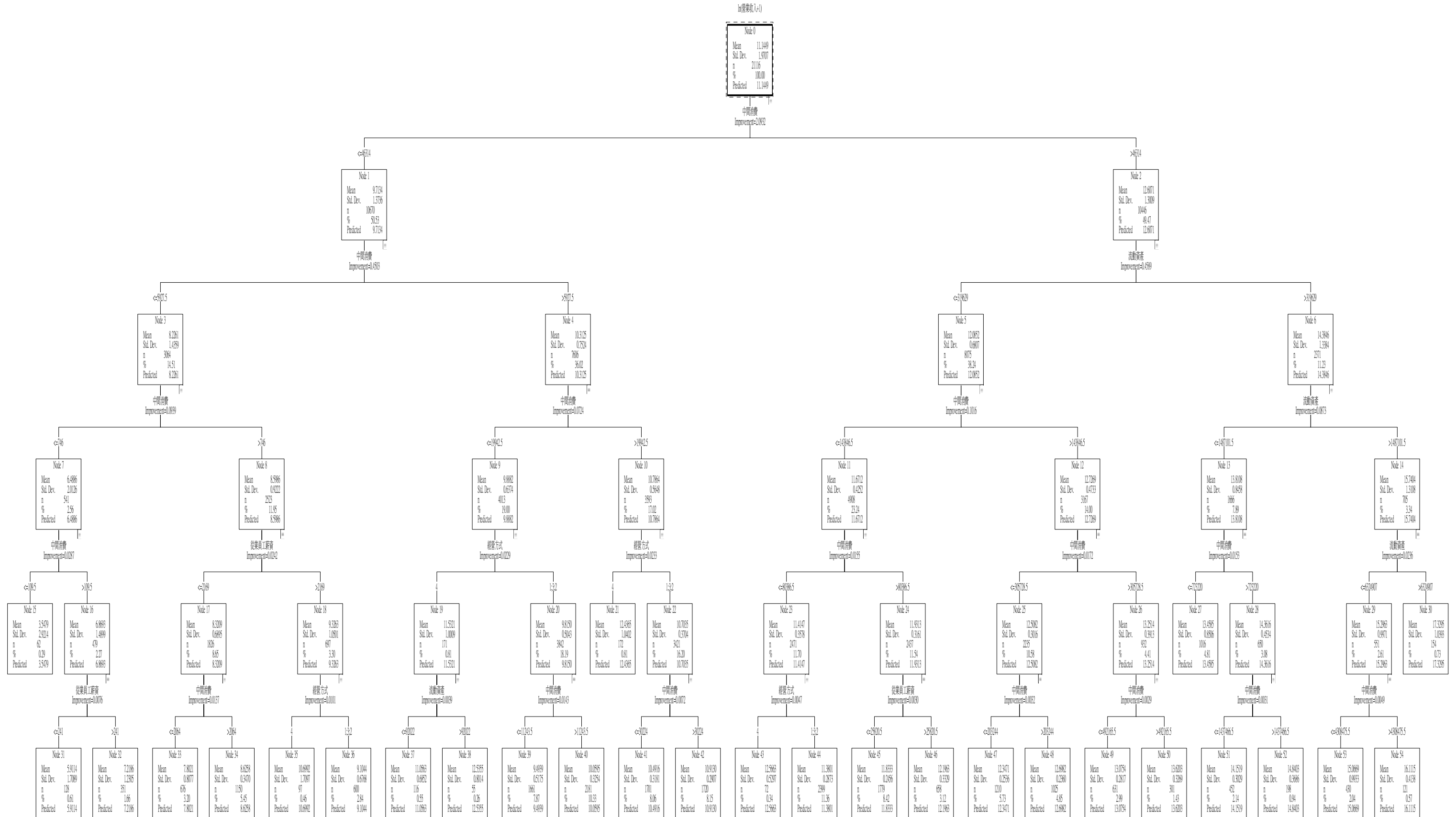


營業收入取0.134912次方 Stem-and-Leaf Plot

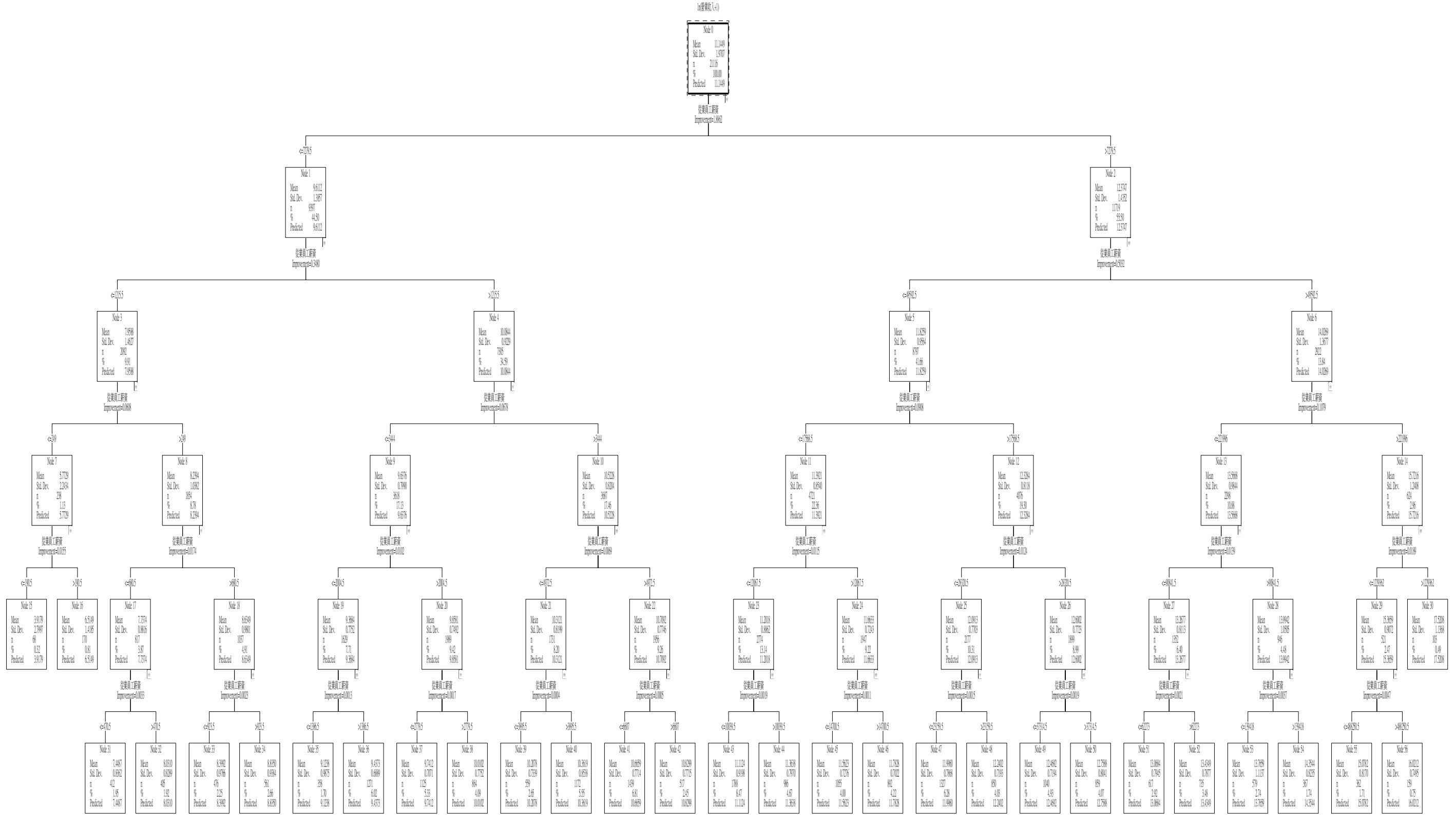
Frequency	Stem & Leaf
2.00	Extremes (≤ 1.86)
1.00	19 . &
2.00	20 . &
6.00	21 . &
4.00	22 . &
11.00	23 . &
171.00	24 . 7799999&&
483.00	25 . 001112222344455556667788999
328.00	26 . 0112334456678899
816.00	27 . 0011122233444445555666677778888999999
742.00	28 . 000111122223333444555666677788899999
719.00	29 . 000111112223333444555666677888999
657.00	30 . 0001111222333344445566677788899
632.00	31 . 000111122233334445556667788999
468.00	32 . 0011122333445566677889
400.00	33 . 001122334455566678899
313.00	34 . 01123344556789
260.00	35 . 01234566789
234.00	36 . 0123456789
178.00	37 . 0123456789
141.00	38 . 012345678&
131.00	39 . 02578&&
125.00	40 . 2458&&
68.00	41 . 0&&&
11.00	42 . &
196.00	Extremes (≥ 4.22)

Stem width: .10
Each leaf: 21 case(s)
& denotes fractional leaves.

附錄 7 CART 樹狀圖



附錄 8 CART 樹狀圖－從業員工薪資



附錄9 地區、營業收入及行業列聯表(4x5x6)

		Cell Counts and Residuals										
地區	營收	行業	Observed		Expected		Residual	Standardized Residual	Adjusted Residual	Deviance		
			Count	%	Count	%						
1	1	1	259	.4%	1580.736	2.2%	-1321.736	-33.610	-37.476	-30.610		
		2	84	.1%	226.732	.3%	-142.732	-9.494	-9.960	-12.916		
		3	106	.1%	1521.770	2.0%	-46.770	-3.747	-3.918	-8.755		
		4	724	1.0%	1118.880	1.5%	-394.880	-11.897	-12.978	-25.106		
		5	122	.2%	161.304	.2%	-39.304	-3.098	-3.241	-8.255		
		6	3341	4.6%	2180.205	3.0%	1160.795	25.240	29.013	53.406		
	2	1	365	.5%	1901.465	2.6%	-1536.465	-35.704	-40.124	-51.711		
		2	123	.2%	272.735	.4%	-149.735	-9.084	-9.607	-13.996		
		3	169	.2%	183.044	.3%	-14.044	-1.039	-1.096	-5.194		
		4	662	.9%	1345.900	1.8%	-683.900	-18.816	-20.688	-30.680		
		5	105	.1%	194.033	.3%	-89.033	-6.400	-6.751	-11.356		
		6	4386	6.0%	2627.566	3.6%	1763.434	35.071	40.631	67.165		
	3	1	2645	3.6%	3351.931	4.6%	-706.931	-12.501	-14.594	-25.398		
		2	464	.6%	480.782	.7%	-16.782	-1.268	-1.344	-5.742		
		3	402	.6%	322.674	.4%	79.326	4.426	4.851	13.294		
		4	2069	2.8%	2375.572	3.3%	-303.572	-6.336	-7.237	-12.802		
		5	341	.5%	342.043	.4%	-1.043	-0.077	-0.822	-3.443		
		6	4635	6.4%	4623.097	6.3%	11.903	.181	.218	4.882		
4	1	3764	5.2%	2216.636	3.0%	1547.364	33.377	37.807	63.135			
	2	454	.6%	317.942	.4%	136.058	7.647	8.154	17.985			
	3	286	.4%	213.364	.3%	72.616	4.978	5.292	12.944			
	4	2819	3.9%	1568.985	2.1%	1250.015	31.903	35.357	57.477			
	5	415	.6%	226.194	.3%	188.806	12.573	13.371	22.444			
	6	998	1.4%	3057.261	4.2%	-2059.261	-38.049	-44.433	-77.271			
5	1	1848	2.5%	842.490	1.2%	1005.510	34.844	38.183	53.981			
	2	216	.3%	120.842	.2%	95.158	8.664	9.225	15.840			
	3	167	.2%	81.102	.1%	85.898	9.543	9.800	15.532			
	4	1246	1.7%	596.333	.8%	649.667	26.713	28.634	42.853			
	5	386	.5%	85.971	.1%	300.029	32.378	33.262	34.050			
	6	231	.3%	1161.991	1.6%	-840.991	-24.870	-28.096	-29.788			
2	1	1	406	.6%	832.447	1.1%	-426.447	-14.865	-15.885	-24.146		
		2	82	.1%	119.402	.2%	-37.402	-3.426	-3.517	-7.850		
		3	59	.1%	80.135	.1%	-21.135	-2.362	-2.420	-6.011		
		4	581	.8%	589.225	.8%	-8.225	-.340	-.358	-4.042		
		5	57	.1%	84.946	.1%	-27.946	-3.034	-3.109	-7.444		
		6	2014	2.8%	1148.139	1.6%	865.861	25.757	28.055	47.578		
	2	1	592	.8%	1001.349	1.4%	-409.349	-13.026	-14.007	-24.946		
		2	104	.1%	143.628	.2%	-39.628	-3.310	-3.413	-8.194		
		3	55	.1%	96.395	.1%	-41.395	-4.219	-4.341	-7.856		
		4	400	.5%	708.778	1.0%	-308.778	-11.655	-12.348	-21.393		
		5	46	.1%	102.181	.1%	-56.181	-5.562	-5.724	-8.569		
		6	2435	3.3%	1381.095	1.9%	1053.905	28.631	31.412	52.551		
3	1	2579	3.5%	1765.193	2.4%	813.807	19.608	21.721	44.222			
	2	272	.4%	253.189	.3%	18.811	1.246	1.314	6.117			
	3	150	.2%	169.926	.2%	-19.926	-1.530	-1.607	-6.114			
	4	1003	1.4%	1249.443	1.7%	-246.443	-7.032	-7.649	-20.993			
	5	91	.1%	180.127	.2%	-89.127	-6.649	-6.982	-11.148			
	6	2278	3.1%	2434.615	3.2%	-156.615	-3.228	-3.461	-7.005			
4	1	2464	3.4%	1167.325	1.6%	1296.675	38.259	41.399	60.676			
	2	169	.2%	167.434	.2%	1.566	.121	.125	1.774			
	3	51	.1%	112.372	.2%	-61.372	-5.794	-5.987	-8.977			
	4	765	1.0%	826.259	1.1%	-61.259	-2.143	-2.283	-10.856			
	5	116	.2%	119.118	.2%	-3.118	-.286	-.291	-3.681			
	6	207	.3%	1610.014	2.2%	-1403.014	-35.358	-39.076	-29.142			
5	1	607	.8%	443.672	.6%	163.328	7.778	8.196	19.507			
	2	30	.0%	63.638	.1%	-33.638	-4.219	-4.288	-6.717			
	3	18	.0%	42.710	.1%	-24.710	-3.782	-3.877	-5.577			
	4	163	.2%	314.041	.4%	-151.041	-8.542	-8.941	-14.621			
	5	26	.0%	45.274	.1%	-19.274	-2.865	-2.908	-5.370			
	6	44	.1%	611.928	.8%	-567.928	-23.055	-24.711	-45.230			
3	1	1	347	.5%	913.669	1.3%	-566.669	-18.866	-20.249	-25.921		
		2	65	.1%	121.052	.2%	-66.052	-5.775	-5.942	-9.488		
		3	87	.1%	87.954	.1%	-.954	-.102	-.105	-1.378		
		4	752	1.0%	646.716	.9%	105.284	4.159	4.397	15.061		
		5	100	.1%	93.234	.1%	6.766	.701	.720	3.743		
		6	2085	2.9%	1260.164	1.7%	824.836	23.439	25.671	45.823		
	2	1	527	.7%	1099.052	1.5%	-572.052	-17.387	-18.782	-27.853		
		2	134	.2%	157.642	.2%	-23.642	-1.885	-1.949	-6.599		
		3	79	.1%	105.800	.1%	-26.800	-2.607	-2.690	-6.794		
		4	564	.8%	77.993	.1%	486.007	13.711	14.438	20.946		
		5	43	.1%	112.151	.2%	-69.151	-6.535	-6.743	-9.800		
		6	2843	3.9%	1515.850	2.1%	1327.150	34.447	38.003	59.798		
	3	1	2103	2.9%	1937.424	2.7%	165.576	3.813	4.246	18.572		
		2	435	.6%	277.893	.4%	157.107	9.442	9.983	19.745		
		3	243	.3%	385.506	.5%	-142.506	-4.364	-4.591	-11.340		
		4	1603	2.2%	1371.252	1.9%	231.648	6.315	6.803	22.949		
		5	134	.2%	197.702	.3%	-63.702	-4.537	-4.785	-10.209		
		6	2698	3.7%	2672.162	3.7%	25.838	.509	.582	7.206		
4	1	1919	2.6%	1281.221	1.8%	637.779	17.976	19.543	39.577			
	2	226	.3%	183.771	.2%	42.229	3.119	3.340	9.669			
	3	93	.1%	123.337	.2%	-30.337	-2.734	-2.834	-7.246			
	4	1028	1.4%	906.877	1.2%	121.123	4.047	4.330	16.055			
	5	122	.2%	130.741	.2%	-8.741	-.765	-.793	-4.109			
	6	260	.4%	1761.304	2.4%	-1507.304	-36.294	-40.336	-51.568			
5	1	665	.9%	486.861	.7%	178.039	8.095	8.565	20.558			
	2	69	.1%	69.847	.1%	-.847	-.101	-.103	-1.298			
	3	30	.0%	46.877	.1%	-16.877	-2.466	-2.505	-5.175			
	4	261	.4%	344.682	.5%	-83.682	-4.518	-4.716	-12.049			
	5	30	.0%	49.691	.1%	-19.691	-2.794	-2.839	-5.503			
	6	62	.1%	671.634	.9%	-609.634	-23.633	-25.467	-47.188			
4	1	1	15	.0%	74.046	.1%	-59.046	-6.865	-6.921	-9.621		
		2	18	.0%	10.621	.0%	7.379	2.264	2.277	4.358		
		3	9	.0%	7.128	.0%	1.872	.701	.704	2.049		
		4	108	.1%	524.111	1.1%	-416.111	-12.681	-13.385	-24.497		
		5	3	.0%	7.556	.0%	-4.556	-1.658	-1.665	-2.354		
		6	237	.3%	102.127	.1%	134.873	13.355	13.867	19.976		
	2	1	24	.0%	89.070	.1%	-65.070	-6.899	-7.127	-9.934		
		2	26	.0%	127.776	.0%	13.224	3.700	3.724	6.079		
		3	5	.0%	8.574	.0%	-3.574	-1.221	-1.221	-2.322		
		4	75	.1%	63.046	.1%	11.954	1.506	1.542	5.103		
		5	2	.0%	9.089	.0%	-7.089	-2.352	-2.364	-4.261		
		6	263	.4%	122.848	.2%	140.152	12.656	13.233	20.010		
	3	1	98	.1%	157.014	.2%	-59.014	-4.715	-4.987	-9.612		
		2	50	.1%	22.521	.0%	27.479	5.791	5.855	8.931		
		3	28	.0%	15.115	.0%	12.885	3.315	3.343	5.876		
		4	171	.2%	111.138	.2%	59.862	5.683	5.916	12.139		
		5	14	.0%	16.022	.0%	-2.022	-.505	-.510	-1.944		
		6	221	.3%	216.559	.3%	4.441	.302	.327	2.996		
4	1	52	.1%	103.833	.1%	-51.833	-5.090	-5.285	-8.481			
	2	30	.0%	14.893	.0%	15.107	3.915	3.944	6.482			
	3	9	.0%	9.996	.0%	-.996	-.315	-.317	-1.374			
	4	67	.1%	73.496	.1%	-6.496	-.758	-.779	-3.521			
	5	16	.0%	10.596	.0%	5.404	1.660	1.670	3.632			
	6	22	.0%	143.211	.2%	-121.211	-10.139	-10.676	-19.079			
5	1	6	.0%	39.465	.1%	-33.465	-5.328	-5.414	-7.474			
	2	5.661	.0%	3.799	.0%	-3.661	-1.539	-1.546	-2.040			
	3	3	.0%	3.799	.0%	-.799	-.410	-.411	-1.190			
	4	3	.0%	4.027	.0%	-1.027	-.512	-.513	-1.329			
	5	7	.0%	54.431	.1%	-47.431	-6.431	-6.573	-8.359			
	6	7	.0%	54.431	.1%	-47.431	-6.431	-6.573	-8.359			

a. Model: Multinomial
b. Design: Constant + 地區 + 行業 + 營收

附錄10 地區、營業收入及行業樣本資料對數線性模型參數估計

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Constant	2.015(a)					
[地區 = 1]	3.758	.369	10.174	.000	3.034	4.482
[地區 = 2]	1.781	.395	4.511	.000	1.007	2.554
[地區 = 3]	2.120	.386	5.487	.000	1.363	2.878
[營收 = 1]	3.455	.371	9.317	.000	2.728	4.182
[營收 = 2]	3.559	.370	9.611	.000	2.833	4.285
[營收 = 3]	3.386	.371	9.119	.000	2.658	4.113
[營收 = 4]	1.099	.422	2.606	.009	.272	1.925
[行業 = 1]	-.143	.536	-.267	.789	-1.193	.907
[行業 = 2]	-1.099	.730	-1.504	.132	-2.530	.333
[行業 = 3]	-.762	.647	-1.177	.239	-2.031	.507
[行業 = 4]	-.310	.561	-.552	.581	-1.410	.790
[行業 = 5]	-.762	.647	-1.177	.239	-2.031	.507
[地區 = 1] * [營收 = 1]	-1.114	.375	-2.967	.003	-1.850	-.378
[地區 = 1] * [營收 = 2]	-.946	.375	-2.524	.012	-1.680	-.211
[地區 = 1] * [營收 = 3]	-.717	.376	-1.908	.056	-1.453	.019
[地區 = 1] * [營收 = 4]	.035	.426	.081	.935	-.801	.871
[地區 = 2] * [營收 = 1]	.357	.401	.892	.372	-.428	1.143
[地區 = 2] * [營收 = 2]	.443	.400	1.108	.268	-.341	1.227
[地區 = 2] * [營收 = 3]	.550	.401	1.372	.170	-.236	1.336
[地區 = 2] * [營收 = 4]	.441	.453	.974	.330	-.447	1.329
[地區 = 3] * [營收 = 1]	.052	.392	.133	.894	-.717	.822
[地區 = 3] * [營收 = 2]	.258	.392	.660	.509	-.509	1.026
[地區 = 3] * [營收 = 3]	.380	.393	.967	.334	-.390	1.149
[地區 = 3] * [營收 = 4]	.329	.445	.740	.459	-.542	1.200
[地區 = 1] * [行業 = 1]	1.892	.539	3.509	.000	.835	2.949
[地區 = 1] * [行業 = 2]	.703	.736	.956	.339	-.738	2.145
[地區 = 1] * [行業 = 3]	.110	.654	.168	.866	-1.172	1.393
[地區 = 1] * [行業 = 4]	1.665	.565	2.948	.003	.558	2.772
[地區 = 1] * [行業 = 5]	.946	.652	1.452	.147	-.331	2.224
[地區 = 2] * [行業 = 1]	2.757	.558	4.941	.000	1.663	3.851
[地區 = 2] * [行業 = 2]	.721	.767	.940	.347	-.783	2.224
[地區 = 2] * [行業 = 3]	-.116	.704	-.164	.870	-1.495	1.264
[地區 = 2] * [行業 = 4]	1.611	.586	2.749	.006	.462	2.761
[地區 = 2] * [行業 = 5]	.244	.692	.352	.725	-1.113	1.601
[地區 = 3] * [行業 = 1]	2.508	.552	4.544	.000	1.427	3.590
[地區 = 3] * [行業 = 2]	1.205	.751	1.605	.109	-.267	2.676
[地區 = 3] * [行業 = 3]	.045	.684	.065	.948	-1.296	1.385
[地區 = 3] * [行業 = 4]	1.741	.579	3.009	.003	.607	2.876
[地區 = 3] * [行業 = 5]	.045	.684	.065	.948	-1.296	1.385
[營收 = 1] * [行業 = 1]	-2.586	.597	-4.335	.000	-3.755	-1.417
[營收 = 1] * [行業 = 2]	-1.454	.769	-1.890	.059	-2.961	.054
[營收 = 1] * [行業 = 3]	-2.457	.727	-3.379	.001	-3.882	-1.032
[營收 = 1] * [行業 = 4]	-.473	.573	-.826	.409	-1.597	.650
[營收 = 1] * [行業 = 5]	-3.455	.842	-4.104	.000	-5.106	-1.805

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
[營收 = 2] * [行業 = 1]	-2.232	.576	-3.875	.000	-3.361	-1.103
[營收 = 2] * [行業 = 2]	-1.198	.758	-1.581	.114	-2.684	.288
[營收 = 2] * [行業 = 3]	-3.107	.778	-3.996	.000	-4.631	-1.583
[營收 = 2] * [行業 = 4]	-.940	.576	-1.631	.103	-2.069	.190
[營收 = 2] * [行業 = 5]	-3.896	.907	-4.295	.000	-5.673	-2.118
[營收 = 3] * [行業 = 1]	-.667	.549	-1.215	.225	-1.744	.410
[營收 = 3] * [行業 = 2]	-.380	.747	-.509	.611	-1.843	1.084
[營收 = 3] * [行業 = 3]	-1.288	.677	-1.902	.057	-2.616	.039
[營收 = 3] * [行業 = 4]	.054	.571	.095	.924	-1.064	1.173
[營收 = 3] * [行業 = 5]	-1.964	.702	-2.799	.005	-3.340	-.589
[營收 = 4] * [行業 = 1]	.990	.592	1.672	.094	-.170	2.151
[營收 = 4] * [行業 = 2]	1.403	.781	1.795	.073	-.129	2.934
[營收 = 4] * [行業 = 3]	-.100	.754	-.133	.894	-1.578	1.378
[營收 = 4] * [行業 = 4]	1.409	.612	2.302	.021	.209	2.608
[營收 = 4] * [行業 = 5]	.452	.724	.624	.532	-.967	1.871
[地區 = 1] * [營收 = 1] * [行業 = 1]	-1.718	.603	-2.849	.004	-2.900	-.536
[地區 = 1] * [營收 = 1] * [行業 = 2]	-1.828	.782	-2.338	.019	-3.361	-.296
[地區 = 1] * [營收 = 1] * [行業 = 3]	-.337	.740	-.456	.648	-1.787	1.113
[地區 = 1] * [營收 = 1] * [行業 = 4]	-2.411	.578	-4.170	.000	-3.544	-1.278
[地區 = 1] * [營收 = 1] * [行業 = 5]	-.035	.850	-.041	.967	-1.702	1.632
[地區 = 1] * [營收 = 2] * [行業 = 1]	-2.002	.582	-3.441	.001	-3.142	-.862
[地區 = 1] * [營收 = 2] * [行業 = 2]	-1.976	.769	-2.571	.010	-3.483	-.470
[地區 = 1] * [營收 = 2] * [行業 = 3]	.506	.787	.642	.521	-1.037	2.049
[地區 = 1] * [營收 = 2] * [行業 = 4]	-2.306	.581	-3.967	.000	-3.445	-1.166
[地區 = 1] * [營收 = 2] * [行業 = 5]	-.016	.916	-.018	.986	-1.811	1.778
[地區 = 1] * [營收 = 3] * [行業 = 1]	-1.643	.553	-2.969	.003	-2.727	-.558
[地區 = 1] * [營收 = 3] * [行業 = 2]	-1.525	.753	-2.024	.043	-3.002	-.049
[地區 = 1] * [營收 = 3] * [行業 = 3]	-.503	.686	-.734	.463	-1.848	.841
[地區 = 1] * [營收 = 3] * [行業 = 4]	-2.216	.575	-3.857	.000	-3.342	-1.090
[地區 = 1] * [營收 = 3] * [行業 = 5]	-.828	.708	-1.170	.242	-2.216	.560
[地區 = 1] * [營收 = 4] * [行業 = 1]	-1.412	.596	-2.369	.018	-2.581	-.244
[地區 = 1] * [營收 = 4] * [行業 = 2]	-1.794	.788	-2.276	.023	-3.340	-.249
[地區 = 1] * [營收 = 4] * [行業 = 3]	-.496	.763	-.651	.515	-1.992	.999
[地區 = 1] * [營收 = 4] * [行業 = 4]	-1.726	.616	-2.801	.005	-2.933	-.518
[地區 = 1] * [營收 = 4] * [行業 = 5]	-1.513	.730	-2.072	.038	-2.944	-.082
[地區 = 2] * [營收 = 1] * [行業 = 1]	-1.628	.619	-2.631	.009	-2.841	-.415
[地區 = 2] * [營收 = 1] * [行業 = 2]	-1.364	.812	-1.679	.093	-2.955	.228
[地區 = 2] * [營收 = 1] * [行業 = 3]	-.188	.789	-.238	.812	-1.734	1.359
[地區 = 2] * [營收 = 1] * [行業 = 4]	-2.071	.599	-3.454	.001	-3.246	-.896
[地區 = 2] * [營收 = 1] * [行業 = 5]	.417	.887	.470	.638	-1.322	2.156
[地區 = 2] * [營收 = 2] * [行業 = 1]	-1.795	.598	-3.000	.003	-2.968	-.622
[地區 = 2] * [營收 = 2] * [行業 = 2]	-1.573	.800	-1.966	.049	-3.141	-.005
[地區 = 2] * [營收 = 2] * [行業 = 3]	.203	.836	.243	.808	-1.436	1.843
[地區 = 2] * [營收 = 2] * [行業 = 4]	-2.167	.603	-3.593	.000	-3.349	-.985
[地區 = 2] * [營收 = 2] * [行業 = 5]	.456	.951	.479	.632	-1.409	2.320
[地區 = 2] * [營收 = 3] * [行業 = 1]	-1.823	.572	-3.188	.001	-2.943	-.702
[地區 = 2] * [營收 = 3] * [行業 = 2]	-1.366	.785	-1.739	.082	-2.906	.173
[地區 = 2] * [營收 = 3] * [行業 = 3]	-.551	.736	-.749	.454	-1.994	.892

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
[地區 = 2] * [營收 = 3] * [行業 = 4]	-2.176	.596	-3.649	.000	-3.344	-1.007
[地區 = 2] * [營收 = 3] * [行業 = 5]	-.732	.751	-.975	.329	-2.205	.740
[地區 = 2] * [營收 = 4] * [行業 = 1]	-1.130	.616	-1.832	.067	-2.338	.079
[地區 = 2] * [營收 = 4] * [行業 = 2]	-1.227	.822	-1.492	.136	-2.839	.385
[地區 = 2] * [營收 = 4] * [行業 = 3]	-.416	.818	-.508	.611	-2.019	1.188
[地區 = 2] * [營收 = 4] * [行業 = 4]	-1.405	.640	-2.196	.028	-2.658	-.151
[地區 = 2] * [營收 = 4] * [行業 = 5]	-.511	.773	-.661	.509	-2.026	1.004
[地區 = 3] * [營收 = 1] * [行業 = 1]	-1.571	.614	-2.560	.010	-2.774	-.368
[地區 = 3] * [營收 = 1] * [行業 = 2]	-2.113	.799	-2.646	.008	-3.678	-.548
[地區 = 3] * [營收 = 1] * [行業 = 3]	.003	.768	.004	.997	-1.501	1.508
[地區 = 3] * [營收 = 1] * [行業 = 4]	-1.977	.592	-3.341	.001	-3.137	-.817
[地區 = 3] * [營收 = 1] * [行業 = 5]	1.140	.876	1.301	.193	-.578	2.858
[地區 = 3] * [營收 = 2] * [行業 = 1]	-1.818	.593	-3.066	.002	-2.980	-.656
[地區 = 3] * [營收 = 2] * [行業 = 2]	-1.959	.783	-2.502	.012	-3.494	-.425
[地區 = 3] * [營收 = 2] * [行業 = 3]	.248	.816	.303	.762	-1.352	1.848
[地區 = 3] * [營收 = 2] * [行業 = 4]	-2.108	.595	-3.543	.000	-3.275	-.942
[地區 = 3] * [營收 = 2] * [行業 = 5]	.433	.946	.458	.647	-1.421	2.287
[地區 = 3] * [營收 = 3] * [行業 = 1]	-1.947	.566	-3.441	.001	-3.056	-.838
[地區 = 3] * [營收 = 3] * [行業 = 2]	-1.550	.769	-2.017	.044	-3.057	-.044
[地區 = 3] * [營收 = 3] * [行業 = 3]	-.400	.715	-.558	.577	-1.802	1.003
[地區 = 3] * [營收 = 3] * [行業 = 4]	-2.006	.588	-3.409	.001	-3.160	-.853
[地區 = 3] * [營收 = 3] * [行業 = 5]	-.317	.741	-.428	.668	-1.770	1.135
[地區 = 3] * [營收 = 4] * [行業 = 1]	-1.359	.610	-2.226	.026	-2.555	-.162
[地區 = 3] * [營收 = 4] * [行業 = 2]	-1.649	.806	-2.046	.041	-3.228	-.070
[地區 = 3] * [營收 = 4] * [行業 = 3]	-.207	.795	-.261	.794	-1.765	1.351
[地區 = 3] * [營收 = 4] * [行業 = 4]	-1.467	.632	-2.322	.020	-2.705	-.229
[地區 = 3] * [營收 = 4] * [行業 = 5]	-.489	.765	-.639	.523	-1.988	1.010

a.Constants are not parameters under the multinomial assumption. Therefore, their standard errors are not calculated.

b.This parameter is set to zero because it is redundant..

c..Model: Multinomial.

d.Design: Constant + 地區 + 營業收入 + 行業 + 地區 * 營業收入 + 地區 * 行業 + 營業收入 * 行業 + 地區 * 營業收入 * 行業