

南華大學科技學院資訊管理學系

碩士論文

Department of Information Management

College of Science and Technology

Nanhua University

Master Thesis

整合不同資料來源之資料方體與分析系統的研究

—以癌症病患中醫證候與氣象因子為例

Research on Data Cube and Analysis System Integrating Different  
Data Sources--Take TCM Syndromes of Cancer Patients and  
Meteorological Factors as an Example

黃信凱

Xin-Kai Huang

指導教授：邱宏彬 博士

Advisor: Hung-Pin Chiu, Ph.D.

中華民國 109 年 7 月

July 2020

南華大學  
科技學院資訊管理學系  
碩士學位論文

整合不同資料來源之資料方體與分析系統的研究  
—以癌症病患中醫證候與氣象因子為例

Research on Data Cube and Analysis System Integrating  
Different Data Sources  
--Take TCM Syndromes of Cancer Patients and Meteorological  
Factors as an Example

研究生：甘信凱

經考試合格特此證明

口試委員：林迺衛  
陳張宗榮  
葉明憲  
邱宏彬

指導教授：邱宏彬  
系主任(所長)：陳心良

口試日期：中華民國 109 年 7 月 2 日

## 南華大學資訊管理學系碩士論文著作財產權同意書

立書人：\_\_\_\_\_黃信凱\_\_\_\_\_之碩士畢業論文

中文題目：

整合不同資料來源之資料方體與分析系統的研究－以癌症病患中醫證候與氣象因子為例

英文題目：

Research on Data Cube and Analysis System Integrating Different Data Sources --Take TCM Syndromes of Cancer Patients and Meteorological Factors as an Example

指導教授： 邱宏彬 博士

學生與指導老師就本篇論文內容及資料其著作財產權歸屬如下：

- 共同享有著作權
- 共同享有著作權，學生願「拋棄」著作財產權
- 學生獨自享有著作財產權

學 生： 黃信凱 (請親自簽名)

指導老師： 邱宏彬 (請親自簽名)

中 華 民 國 109 年 7 月 13 日

## 誌 謝

感謝我的指導教授—邱宏彬老師，老師教導許多不同的資訊技術，同時培養指導與教學的能力，在研究所期間讓我擔任助教，在課堂後指導與教學大學生，學習如何將理解的知識，轉化成指導人的方式，同時也重新理解內容，促使我收穫良多。

感謝陳信良主任與尤國任老師，在學期間給我很多未來方向與建議，讓我對於自身更加理解，也更能面對未來的挑戰。

感謝資管系助青翰及燕琪，前者鼓勵我唸研究所並予以協助，後者讓我成為系辦的行政助理，了解許多行政與教學的流程，並且學會如何與其他處室跟院系助理溝通，學會待人處事的道理。

感謝我的學長姐們，育嘉、俊億、家熒、紀延，不論是課業、系務或未來方向上，他們都極力地幫助我。

感謝我的學弟妹們，家綸、憶臻、宜瑾，在系務與助教上幫我分擔很多事情，也在撰寫論文期間消解我的許多心理壓力。

感謝口試委員，陳張宗榮老師、林迺衛老師與葉明憲醫師，因為有口委們的建議，讓我理解論文的不足，能將論文修改至更好。

最後感謝家人及朋友不斷支持我，讓我能夠順利地走完研究所這條路，謝謝你們，也謝謝各位幫助過我的人。

黃信凱 謹誌

民國一零九年七月

# 整合不同資料來源之資料方體與分析系統的研究 — 以癌症病患中醫證候與氣象因子為例

學生：黃信凱

指導教授：邱宏彬

南 華 大 學 資 訊 管 理 學 系 碩 士 班

## 摘 要

鑑於近幾十年來氣象逐漸異常，各行各業在決定方針與策略時，開始顧慮氣象因素，進而發展出氣象經濟與氣象醫學等等與氣象有關之事物。本研究利用資料倉儲之萃取、轉換與載入之概念，以癌症病患中醫證候與氣象因子為例，將此兩種資料進行整合，建構出中醫證候與氣象因子資料方體，從資料方體取出不同時間維度資料進行敘述性統計，再從資料方體取出資料探勘所需之資料，使用 RapidMiner 資料探勘軟體進行決策樹分析，找出適合不同型態之輸入變數模型，最後上述兩種方式，找出在12、1月與10、11月時，平均氣壓最影響數值型態與最低氣溫最影響類別型態，與其他氣象因子在不同規則下皆會影響肝火上炎證候。

中文關鍵字：中醫證候、氣象因子、資料方體、決策樹分析

Research on Data Cube and Analysis System Integrating  
Different Data Sources  
--Take TCM Syndromes of Cancer Patients and Meteorological  
Factors as an Example

Student: Xin-Kai Huang

Advisor: Hung-Pin CHIU, Ph.D.

Department of Information Management  
Nanhua University  
Master Thesis

ABSTRACT

In view of the gradual abnormality of meteorology in recent decades, various industries began to consider meteorological factors when deciding policies and strategies, and then developed meteorological economy and meteorological medicine and other meteorological-related things. In this study, the concepts of extraction, transform and loading of data warehouses are used to take the TCM syndromes of cancer patients and meteorological factors as examples, and the two data are integrated to construct a cube of TCM syndromes and meteorological factors. From the data cube extracts data from different time dimensions for descriptive statistics, and then extracts the data required for data mining from the data cube, and uses RapidMiner data mining software for decision tree analysis to find input variable models suitable for different types. Finally, the above two methods find out that in December, January, October, and November, the average air pressure most affects the numerical type and the minimum temperature most affects the nominal type, and other

meteorological factors under different rules will affect syndromes of Flaring-up of Liver Fire.

**KeyWords:** TCM syndrome, Meteorological factor, Data cube, Decision tree analysis



# 目錄

著作財產權同意書.....	I
誌謝.....	II
摘要.....	III
ABSTRACT .....	IV
目錄.....	VI
圖目錄.....	VIII
表目錄.....	XII
第一章 緒論 .....	1
1.1. 研究動機.....	1
1.2. 研究目的.....	2
1.3. 研究架構.....	3
第二章 文獻探討.....	4
2.1. 中醫證候與辨證.....	4
2.2. 氣象醫學.....	4
2.3. 資料探勘.....	5
2.4. 資料倉儲(Data Warehouse) .....	10
第三章 研究方法.....	14
3.1. 研究流程.....	14
3.2. 中醫門診資料庫.....	17
3.3. 氣象因子資料庫.....	17
3.4. 資料篩選與預處理.....	18
3.5. 資料分群.....	21
3.6. 資料方體.....	24



第四章資料分析.....	27
4.1. 敘述性統計.....	28
4.2. RapidMiner 資料探勘分析.....	63
4.3. 分析結果討論.....	77
第五章結論與未來發展.....	95
5.1. 結論.....	95
5.2. 未來發展.....	96
參考文獻.....	97



## 圖目錄

圖 1 研究目的圖 .....	2
圖 2 資料庫知識挖掘流程圖(KDD Process) .....	6
圖 3 決策樹圖 .....	8
圖 4 星狀綱要結構圖 .....	11
圖 5 某公司產品銷售資料方體圖 .....	12
圖 6 某公司產品銷售資料方體之下鑽與上卷圖 .....	12
圖 7 某公司產品銷售資料方體切片圖 .....	13
圖 8 某公司產品銷售資料方體切塊圖 .....	13
圖 9 研究流程圖 .....	15
圖 10 氣象因子區間與累積證候人數區間 .....	18
圖 11 中醫標準症狀對應證候 .....	19
圖 12 RapidMiner 資料探勘軟體之 k-Means 分群法流程 .....	21
圖 13 RapidMiner 資料探勘軟體之 k-Means 分群法結果-以最大日降水量為例 .....	21
圖 14 RapidMiner 資料探勘軟體之 Filter Expamples 應用圖 .....	22
圖 15 RapidMiner 資料探勘軟體之 k-Means 分群法結果(處理過後) .....	23
圖 16 資料方體圖-以中醫證候與氣象因子為例 .....	25
圖 17 中醫證候與氣象因子資料方體之星狀綱要結構 .....	26
圖 18 中醫證候維度與時間維度(每年各月)之資料方體切片 .....	28
圖 19 中醫證候(虛證)與時間維度(2010 年各月)之資料切片結果 .....	29
圖 20 中醫證候(實證)與時間維度(2010 年各月)之資料切片結果 .....	29
圖 21 中醫證候(虛證)與時間維度(2011 年各月)之資料切片結果 .....	30
圖 22 中醫證候(實證)與時間維度(2011 年各月)之資料切片結果 .....	30
圖 23 中醫證候(虛證)與時間維度(2012 年各月)之資料切片結果 .....	31
圖 24 中醫證候(實證)與時間維度(2012 年各月)之資料切片結果 .....	31
圖 25 中醫證候(虛證)與時間維度(2013 年各月)之資料切片結果 .....	32
圖 26 中醫證候(實證)與時間維度(2013 年各月)之資料切片結果 .....	32
圖 27 心陰虛每年各月就診人數百分比折線圖 .....	33
圖 28 腎不納氣每年各月就診人數百分比折線圖 .....	33
圖 29 心脈痺阻(寒凝)每年各月就診人數百分比折線圖 .....	34
圖 30 心脈痺阻(氣滯)每年各月就診人數百分比折線圖 .....	34
圖 31 肝火上炎每年各月就診人數百分比折線圖 .....	35
圖 32 肝陽上亢每年各月就診人數百分比折線圖 .....	35
圖 33 中醫證候維度與時間維度(特定月份)之二維度資料方體切割 .....	36
圖 34 中醫證候(虛證)與時間維度(特定月份)之資料切片結果 .....	37
圖 35 中醫證候(實證)與時間維度(特定月份)之資料切片結果 .....	37
圖 36 心陽虛特定月份就診人數百分比長條圖 .....	38

圖 37	腎不納氣特定月份就診人數百分比長條圖.....	38
圖 38	心脈痺阻(寒凝)特定月份就診人數百分比長條圖.....	39
圖 39	心脈痺阻(氣滯)特定月份就診人數百分比長條圖.....	39
圖 40	肝火上炎特定月份就診人數百分比長條圖.....	40
圖 41	肝陽上亢特定月份就診人數百分比長條圖.....	40
圖 42	中醫證候維度與時間維度(每年各季)之二維度資料方體切割.	41
圖 43	中醫證候(虛證)與時間維度(每年各季)之資料切片結果.....	42
圖 44	中醫證候(實證)與時間維度(每年各季)之資料切片結果.....	42
圖 45	心陰虛每年各季就診人數百分比長條圖.....	43
圖 46	腎不納氣每年各季就診人數百分比長條圖.....	43
圖 47	心脈痺阻(寒凝)每年各季就診人數百分比長條圖.....	44
圖 48	心脈痺阻(氣滯)每年各季就診人數百分比長條圖.....	44
圖 49	肝火上炎每年各季就診人數百分比長條圖.....	45
圖 50	肝陽上亢每年各季就診人數百分比長條圖.....	45
圖 51	中醫證候維度與時間維度(特定季節)切割.....	46
圖 52	中醫證候(虛證)與時間維度(特定季節)之資料切片結果.....	47
圖 53	中醫證候(實證)與時間維度(特定季節)之資料切片結果.....	47
圖 54	腎不納氣特定季節就診人數百分比長條圖.....	48
圖 55	暑襲衛表特定季節就診人數百分比長條圖.....	48
圖 56	心脈痺阻(寒凝)特定季節就診人數百分比長條圖.....	49
圖 57	心脈痺阻(氣滯)特定季節就診人數百分比長條圖.....	49
圖 58	肝火上炎特定季節就診人數百分比長條圖.....	50
圖 59	肝陽上亢特定季節就診人數百分比長條圖.....	50
圖 60	中醫證候維度與時間維度(每年六氣)之二維度資料方體切割.	51
圖 61	中醫證候(虛證)與時間維度(2010年六氣)之資料切片結果....	52
圖 62	中醫證候(實證)與時間維度(每年六氣)之資料切片結果.....	52
圖 63	中醫證候(虛證)與時間維度(2011年六氣)之資料切片結果....	53
圖 64	中醫證候(實證)與時間維度(2011年六氣)之資料切片結果....	53
圖 65	中醫證候(虛證)與時間維度(2012年六氣)之資料切片結果....	54
圖 66	中醫證候(實證)與時間維度(2012年六氣)之資料切片結果....	54
圖 67	腎氣不固每年六氣就診人數長條圖.....	55
圖 68	寒襲衛表每年六氣就診人數長條圖.....	55
圖 69	心脈痺阻(寒凝)每年六氣就診人數長條圖.....	56
圖 70	心脈痺阻(氣滯)每年六氣就診人數長條圖.....	56
圖 71	肝火上炎每年六氣就診人數長條圖.....	57
圖 72	肝陽上亢每年六氣就診人數長條圖.....	57
圖 73	中醫證候維度與時間維度(特定六氣)切割.....	58
圖 74	中醫證候(虛證)與時間維度(特定六氣)之資料切片結果.....	59
圖 75	中醫證候(實證)與時間維度(特定六氣)之資料切片結果.....	59

圖 76	心陰虛特定六氣就診人數長條圖.....	60
圖 77	寒襲衛表特定六氣就診人數長條圖.....	60
圖 78	心脈痺阻(寒凝)特定六氣就診人數長條圖.....	61
圖 79	心脈痺阻(氣滯)特定六氣就診人數長條圖.....	61
圖 80	肝火上炎特定六氣就診人數長條圖.....	62
圖 81	肝陽上亢特定六氣就診人數長條圖.....	62
圖 82	資料探勘分析圖—以肝火上炎為例.....	63
圖 83	RapidMiner 資料探勘軟體之決策樹分析流程.....	64
圖 84	gini_index 與訓練資料集為 0.6 之決策樹不剪枝分析結果....	65
圖 85	gini_index 與訓練資料集為 0.7 之決策樹不剪枝分析結果....	65
圖 86	gini_index 與訓練資料集為 0.8 之決策樹不剪枝分析結果....	65
圖 87	gini_index 與訓練資料集為 0.6 之決策樹事前剪枝分析結果	66
圖 88	gini_index 與訓練資料集為 0.7 之決策樹事前剪枝分析結果	66
圖 89	gini_index 與訓練資料集為 0.8 之決策樹事前剪枝分析結果	66
圖 90	gini_index 與訓練資料集為 0.6 之決策樹事後剪枝分析結果	67
圖 91	gini_index 與訓練資料集為 0.7 之決策樹事後剪枝分析結果	67
圖 92	gini_index 與訓練資料集為 0.8 之決策樹事後剪枝分析結果	67
圖 93	gain_ratio 與訓練資料集為 0.6 之決策樹不剪枝分析結果....	68
圖 94	gain_ratio 與訓練資料集為 0.7 之決策樹不剪枝分析結果....	68
圖 95	gain_ratio 與訓練資料集為 0.8 之決策樹不剪枝分析結果....	68
圖 96	gain_ratio 與訓練資料集為 0.6 之決策樹事前剪枝分析結果	69
圖 97	gain_ratio 與訓練資料集為 0.7 之決策樹事前剪枝分析結果	69
圖 98	gain_ratio 與訓練資料集為 0.8 之決策樹事前剪枝分析結果	69
圖 99	gain_ratio 與訓練資料集為 0.6 之決策樹事後剪枝分析結果	70
圖 100	gain_ratio 與訓練資料集為 0.7 之決策樹事後剪枝分析結果 .....	70
圖 101	gain_ratio 與訓練資料集為 0.8 之決策樹事後剪枝分析結果 .....	70
圖 102	gini_index 與訓練資料集為 0.6 之決策樹不剪枝分析結果..	71
圖 103	gini_index 與訓練資料集為 0.7 之決策樹不剪枝分析結果..	71
圖 104	gini_index 與訓練資料集為 0.8 之決策樹不剪枝分析結果..	71
圖 105	gini_index 與訓練資料集為 0.6 之決策樹事前剪枝分析結果 .....	72
圖 106	gini_index 與訓練資料集為 0.7 之決策樹事前剪枝分析結果 .....	72
圖 107	gini_index 與訓練資料集為 0.8 之決策樹事前剪枝分析結果 .....	72
圖 108	gini_index 與訓練資料集為 0.6 之決策樹事後剪枝分析結果 .....	73

圖 109 gini_index 與訓練資料集為 0.7 之決策樹事後剪枝分析結果	73
圖 110 gini_index 與訓練資料集為 0.8 之決策樹事後剪枝分析結果	73
圖 111 gain_ratio 與訓練資料集為 0.6 之決策樹不剪枝分析結果..	74
圖 112 gain_ratio 與訓練資料集為 0.7 之決策樹不剪枝分析結果..	74
圖 113 gain_ratio 與訓練資料集為 0.8 之決策樹不剪枝分析結果..	74
圖 114 gain_ratio 與訓練資料集為 0.6 之決策樹事前剪枝分析結果	75
圖 115 gain_ratio 與訓練資料集為 0.7 之決策樹事前剪枝分析結果	75
圖 116 gain_ratio 與訓練資料集為 0.8 之決策樹事前剪枝分析結果	75
圖 117 gain_ratio 與訓練資料集為 0.6 之決策樹事後剪枝分析結果	76
圖 118 gain_ratio 與訓練資料集為 0.7 之決策樹事後剪枝分析結果	76
圖 119 gain_ratio 與訓練資料集為 0.8 之決策樹事前剪枝分析結果	76
圖 120 數值事後 gain_ratio 模型之決策樹樹狀圖.....	82
圖 121 數值事前 gini_index 模型之決策樹樹狀圖.....	85

## 表目錄

表 一 中醫門診資料欄位篩選表 .....	18
表 二 氣象因子篩選表 .....	20
表 三 k-Means 分群法處理前後範圍差異表 .....	23
表 四 資料探勘分析之資料 .....	24
表 五 六氣時段表 .....	51
表 六 不同維度切割之敘述性統計表(月份) .....	77
表 七 不同維度切割之敘述性統計表(季節) .....	78
表 八 不同維度切割之敘述性統計表(六氣) .....	79
表 九 肝火上炎之不同時間維度統計表 .....	80
表 十 不同輸入變數型態與剪枝之決策樹分析結果 .....	81
表 十一 數值事後 gain_ratio 模型之決策樹規則描述 .....	83
表 十二 數值事後 gain_ratio 模型之氣象因子影響肝火上炎類別表 .....	84
表 十三 數值事前 gini_index 模型之決策樹規則描述 .....	85
表 十四 類別事後 gini_index 模型之規則描述(肝火上炎類別低) .....	86
表 十五 類別事後 gini_index 模型之規則描述(肝火上炎類別中) .....	87
表 十六 類別事後 gini_index 模型之規則描述(肝火上炎類別高) .....	88
表 十七 類別事後 gini_index 模型之氣象因子影響肝火上炎類別表 .....	88
表 十八 類別事前 gini_index 模型之規則描述 .....	89
表 十九 類別事前 gini_index 模型之氣象因子影響肝火上炎類別表 .....	89
表 二十 資料探勘方法之比較表 .....	93
表 二十一 不同論文之研究結果 .....	93

# 第一章 緒論

## 1.1. 研究動機

鑑於近幾十年來氣候逐漸異常，導致各行各業進行策略及決定方針時，開始需要顧慮氣象因素，避免遺漏由氣象所造成之問題，也希冀透過研究氣象，發現與氣象有關之商機與病因。

以統一超商為例，其在近幾十年來推出「二代 POS(point of sales)服務情報系統」，藉由結合分析氣象因素得出「七五三感冒指數」，找出人們處在不同氣象現象時，所購買的物品與行為模式會有所不同，在進貨與上架商品等等行為都有一定的依據與指標，可稱之為「氣象經濟」(楊蕙菁，2004)。

除了上述的氣象經濟，近年來氣象醫學也逐漸受到重視，自古以來，中醫學說即有「外感六淫」之說，「外感六淫」則是古時人類遵循四季規律變化而活，病因上之六氣分別為風、寒、暑、濕、燥、火，其為自然界四季氣候的正常變化，但六氣若有大幅異常變化，導致人體無法適應，則人體易出現病狀，對於這種情形，中醫稱其為外感六淫。透過上述古人所研究之自然氣候變化對疾病的影響，了解氣象對人體健康與疾病的相關性，進一步預防疾病發生，屬於氣象醫學與預防醫學之先驅。

近年來資訊科技發展快速，資料相關技術迅速融入人們生活，人們對其需求量與日俱增，資料相關技術成為目前人們不可或缺的技术。現代資料相關技術如雨後春筍般蓬勃發展，許多行業使用資料相關技術進行各式各樣活動。

李義偉(2004)提到耗費較高之資源須從資料倉儲查詢資料，若預先建立起資料方體(Data Cube)的方法，能降低所耗費之資源。代表資料方體搭配資料倉儲使用，可以達到提升效能之成效。

陳宗仁(2019)透過使用資料探勘技術的 Pearson 關聯係數與單因子變異數分析，研究某月擁有特定證候的雲嘉地區癌症病患人數與該月前一段時間地區的氣候變化，讓雲嘉地區醫療人員藉由該研究了解不同氣候因子等級下中醫證候的變化與病患的證候特質，因而達到預防與判斷的效果。

基於上述，中醫在早期即有氣象醫學與預防醫學之概念，而現代資料技術發達，眾多研究者利用資料探勘、資料方體與資料倉儲進行研究，本研究希冀使用資料倉儲概念建構資料方體並以癌症病患中醫證候與氣象因子為例，有系統性地整合癌症病患中醫證候與氣象因子的相關資料並進行資料探勘分析，找出癌症病患中醫證候與氣象因子之間的規則，藉此完成一個具有系統性流程與分析的研究。

## 1.2. 研究目的

總合上述研究動機，本研究之研究目的如下列四點與(圖 1)所示：

1. 以中醫與氣象因子資料為例，使用資料倉儲之萃取、轉換與載入概念，建立不同資料來源之資料方體，利用資料方體維度可變換之特性，供後續分析對資料進行不同維度切割與使用。
2. 從資料方體中切割不同時間維度之資料，進行敘述性統計。
3. 使用 RapidMiner 資料探勘軟體進行決策樹分析，找出適合不同型態之輸入變數模型。
4. 分析不同型態之輸入變數模型所找出的決策樹規則，並進行綜合討論。

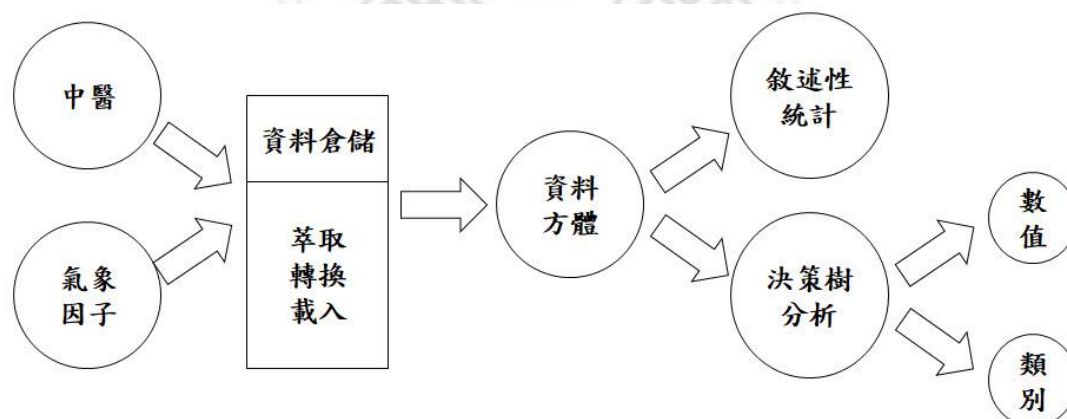


圖 1 研究目的圖



### 1.3. 研究架構

本研究之架構一共分為五大章，藉由將架構分為五大章，將研究動機與理念完整闡述，且將研究過程與結論完整呈現，每章之內容敘述如下：

#### 1. 第一章 緒論

從氣象因子對於現代各行各業之影響做為開端，進而理解氣象醫學之理念，帶入中醫預防醫學之觀念與案例，最後結合資料倉儲與中醫資料探勘之案例，針對不同資料來源之資料方體與分析系統之研究動機目的與架構進行闡述與理解。

#### 2. 第二章 文獻探討

彙整中醫證候與辨證、氣象醫學、資料探勘與資料倉儲之文獻進行引用與定義，探討相關概念，做為本研究之相關依據。

#### 3. 第三章 研究方法

闡述本研究之研究流程，並對中醫與氣象因子資料進行資料蒐集，對資料進行篩選、辨證，遺漏值處理與資料分群，建立具有資料倉儲概念之資料方體。

#### 4. 第四章 資料分析

從資料方體切割出不同資料，進行六種敘述性統計，並利用決策樹分析找出適合不同型態之輸入變數模型，並分析模型之樹狀圖與規則。

#### 5. 第五章 結論與未來發展

對整個論文進行總結，並寫出未來可供後續團隊進行之建議與發展。

## 第二章 文獻探討

本章將彙整中醫證候與辨證、氣象醫學、資料探勘與資料倉儲等文獻進行探討與研究，藉由探討這些文獻驗證本研究之基礎架構，達到有憑有據之合理探討，以建構後續系統。

### 2.1. 中醫證候與辨證

中醫已發展數千年歷史，診斷疾病的基礎方法可以分為四種「望、聞、問、切」合稱為四診，透過四診取得病患之證候，中醫師藉由四診判斷病患該使用何種處方用藥，證候可代表病患不同程度的健康狀況，只要有些許症狀差異，就能夠發現病患健康狀態的變化，判斷可能歸屬的證候並且針對不同的證候使用有效的治療方法(陳世波、劉保延、王永炎、周雪忠、倪青、何麗云，2009;陳宗仁，2019)。

「辨證」為臨床診斷的初步方法，主要是針對病況在特定時間擬定適當療法進行治療，以達到良好的治療效果，因此辨證是必需的步驟。臨床診斷是最能完整了解病情的過程，其產生臨床病例報告主要是為了讓未參與臨床診斷的中醫師，能透過病例報告快速了解在診斷期間的構想，而後其也有利於中醫師群之間的互相學習與溝通。

### 2.2. 氣象醫學

「六氣學說」之說，為黃帝內經整理先人所累積的豐富醫療經驗，以此為基礎形成系統化之理論，對於氣象與人體之間擁有一套推算規則，其之六氣為風、寒、暑、溼、燥、火(胡建勳、劉堯欽，2008)，此為氣象醫學的範例之一，而有許多研究者進行著氣象醫學的研究。

從古至今中醫十分重視氣象對健康的影響，並認為風、寒、暑、濕、燥、火等六氣之氣象因子會對身體產生相對應的證候。因此使用資料探勘技術的 Pearson 關聯係數與單因子變異數分析，研究某月擁有特定證候的雲嘉地區癌症病患人數與該月前一段時間地區的氣候變化，可提供雲嘉地區醫療人員藉由該研究了解不同氣候因子等級下中醫證候的變化與病患的證候特質。其研究結果顯示心脈痺阻(氣滯)、肝火上炎與肝陽上亢等證候和均溫有正相關，肝火上炎證候和氣壓有負相關，及心脈痺阻(寒凝)證候和平均風

速有負相關(陳宗仁, 2019)。

不同的氣象對傳染病流行的影響程度不同,由此可知傳染病的流行與氣象變化有明顯的關聯性,通過觀察氣象變化可預知傳染病的流行情況。傳染病發病存在著周期變動規律,運氣學可通過預測氣象變化,從而預測傳染病的發生,為北京地區傳染病預測和防控提供指導方案(劉宏偉, 2011)。

氣象因子影響成人社區獲得性肺炎中醫證型的發生,特別是對風熱犯肺證與痰熱壅肺證。在廣東省中醫院的成人社區獲得性肺炎中,風熱犯肺證的發病人數與當月平均風速、相對濕度、平均氣溫及前一個月平均相對濕度等有關係;痰熱壅肺證的發病人數與前一個月日照時數和當月平均相對濕度有關係。因此,成人社區獲得性肺炎的預防和控制應更多考慮當地的氣候變化,相關公共衛生政策的制定應考慮當地氣候變化對疾病的影響(黃晶一, 2015)。

藉由以上的氣象醫學研究可以發現,若研究成果有一定成效時,是能影響到決策與方針,促使公共政策有更好的依據與方案。

### 2.3. 資料探勘

資料探勘是近年來使用相當廣泛,是一種用來將大量資料中所隱藏的資訊挖掘出來的一個技術,從已存在的資料挖掘出不明確、未知及潛在可能有用資訊,進而發現連專家都不了解的新關係,視為資料庫知識挖掘(Knowledge Discovery in Database, KDD)的一個過程(Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P., 1996; Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J., 1992; Gruppe, F. H., & Owrang, M. M., 1995)。

其使用很多統計分析與模式建立的方法,從資料中尋找出有用的樣式 (Patterns)、規則(Rules)或關聯性(Relationships)(簡睿志, 2011),從(圖 2)可了解 KDD 的過程。

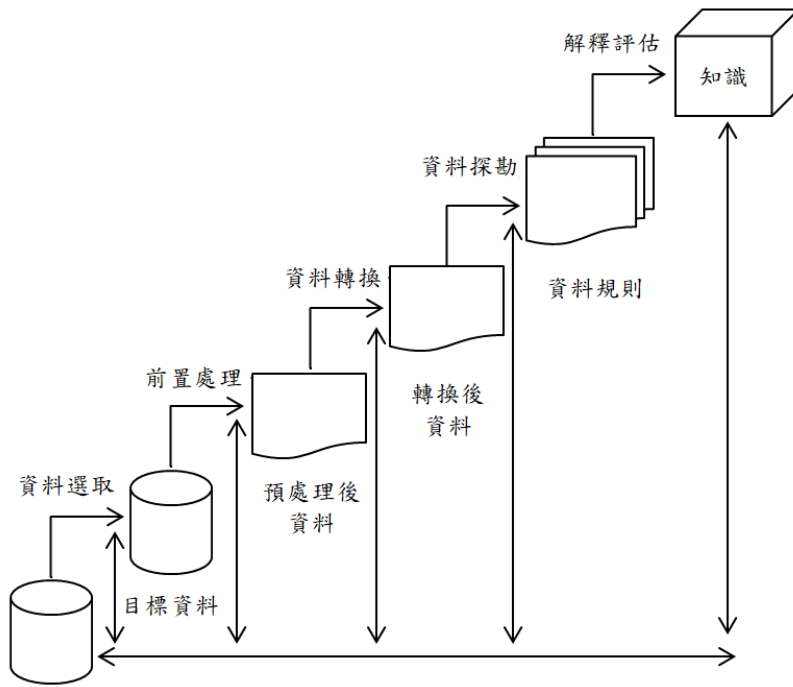


圖 2 資料庫知識挖掘流程圖(KDD Process)

1. 資料選取(Selection):

從資料庫或資料倉儲中選出適合的資料。

2. 前置處理(Pre-processing):

將目標資料(Target Data)進行前置處理，將不合適之資料進行修正與刪除等等動作。

3. 資料轉換(Trans-formation):

將預處理後的資料(Preprocessed Data)進行轉換，轉換成適合的資料格式。

4. 資料探勘(Data Mining):

將轉換完成的資料(Transformed Data)進行探勘分析，找出相關的資料規則。

5. 解釋評估(Interpretation/Evaluation):

將找到的資料規則(Patterns)進行評估與解釋，將規則轉換成為知識(Knowledge)。

### 2.3.1. 決策樹(Decision Tree)

決策樹是在資料探勘的研究領域中常用的一種樹狀結構方法，學習過程採用監督式學習，以訓練資料集建立模型，再以測試資料集驗證模型，建構出具有樹狀之結構，其對特定之原始資料進行分類與歸納，藉此建立合適之決策或預測分析相關判斷規則(陳有忠，2008;林宇恆，2016)，得出具有階層關係之結果。

決策樹是透過「亂度」(Entropy)為分岔概念，所謂的亂度是指屬性分佈的分散程度，亂度若過高，則代表越找不到規則，決策樹藉由降低亂度，亂度最低者即是最佳的樹狀結構(巫美鳳，2015)。

決策樹將資料先行切割，分為訓練資料集(Training Data)與測試資料集(Testing Data)，再依據不同演算法規則得出不同分類與預測，若資料集皆以分類至同類別、該資料集無法得出新節點或該資料集已無能處理之資料，則決策樹就會停止生長(葉子維，2018)，而其結果猶如樹枝般階層呈現，再對其進行樹枝修剪，最終得出之結果相較其他資料探勘方法，更易於為人所了解。

訓練資料集(Training set)是用以訓練模型之資料集，而測試資料集(Testing set)則是用以驗證模型訓練成果之資料集(鄧莉雅，2015)。例如:若將模型視為一位尚未訓練之選手，訓練資料集就可視為教練教導選手之內容，測試資料集則可視為驗證選手訓練成果之大賽。

決策樹是一種樹狀結構圖，從目標變數與其餘變數間找出最好的關係進行分叉，起始節點為根結點，一路依據分類準則分類，中間節點為內部節點，至最後的節點為葉節點，如(圖 3)。

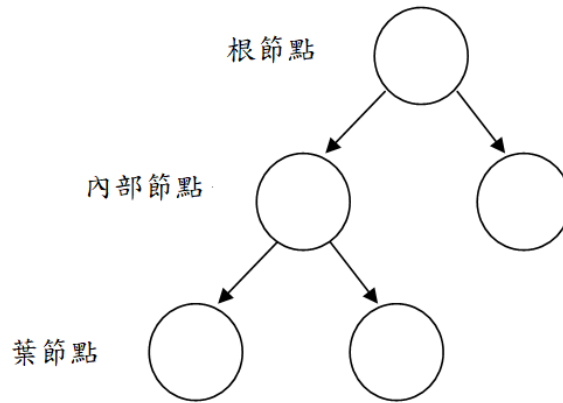


圖 3 決策樹圖

大部份之決策樹演算法能同時處理連續型資料和離散型資料(許政文, 2019), 且能清楚表達目標變數與其餘變數間之階層關係, 使其分類結果能被簡單解釋。

決策樹的生長過程是將資料進行分割, 每次分割所面對問題也有所不同, 也對應著不同節點, 對每個分割都要求分成最大值之間的「差異」, 而這些差異的不同, 也有不同演算法, 其中較為常見的有 ID3、C4.5、CART。

### 2.3.1.1. ID3 演算法

ID3 是 1986 年由 Ross Quinlan 所發明的決策樹演算法, 採用「信息增益 (Information Gain)」, 信息增益為「測試前的信息量」減去「測試後的信息量」(陳有忠, 2008), 在 ID3 演算法中, 信息增益越大, 則優先選擇成為節點, 以此為循環至全部資料計算結束。

ID3 之分類準則較為簡易, 以最大信息增益為選擇條件, 有部份人認為此分類準則過於單純, 會因為某個屬性之數值較多, 認定該屬性有較大解釋力, 且 ID3 演算法僅能處理離散型資料, 也很容易產生過度擬合現象, 故 Quinlan 於 1993 年提出了 C4.5 演算法。

### 2.3.1.2.C4.5 演算法

C4.5 是 1993 年由 Ross Quinlan 所發明的決策樹演算法，從 ID3 演算法改良而來，繼承其優點，並採用「信息增益率(Information Gain Ratio)」，先從劃分屬性中找出高於水平的屬性，再從中選出最高的屬性，改善 ID3 演算法的缺點，可處理連續型資料、克服信息增益會選擇取值多的屬性之問題且可對決策樹進行事後剪枝(顏利憲，2013)。

「信息增益率(Information Gain Ratio)」，是將信息增益除以分割資訊值(Split Information)，因 ID3 產生之決策樹容易過大且單一，藉由分割資訊值(Split Information)修正內部節點過多之問題(林宇恆，2016;鄧莉雅，2015)。

C4.5 演算法所產生之規則易於理解，準確率較高，但其在樹生長過程中，需多次對資訊進行掃描與排序為其缺點。

### 2.3.1.3. CART 演算法

分類回歸樹(Classification And Regression Tree, CART)是 1984 年由 Breiman 所提出，為機械學習的方法之一，其優點為可處理繁雜與多變數的資料、同時處理連續和離散變數，以及離群的資料不會影響其演算(鄧雅莉，2015;吳芮妍，2019)，故在使用面上十分廣闊。

藉由遞迴(Recursive)方式進行二元切割分類，直到資料皆以歸類完成或無法歸類，進而產生一連串規則，生成的樹內部節點代表測試屬性，樹葉節點則表示類別與決策之分佈情形(鄧雅莉，2015)。依照目標變數的不同，可分為以類別變數為主的分類樹(Classification Tree)與以連續變數為主的迴歸樹(Regression Tree)。

CART 的分類準則採用的是 Gini 係數作為屬性的分散度，若 Gini 係數越低，代表節點的資料偏屬於同分類，而 CART 就是透過選擇較低 Gini 係數達到節點的分類純度提升，使節點中資料有較高的同質性，按照此分則生長出決策樹，再進行剪枝。

## 2.4. 資料倉儲(Data Warehouse)

資料倉儲可追溯至 1960 年代，當時已利用大型主機進行產生定期報表；1970 年代，高階主管資訊系統出現在市場上；1980 年代，已有公司開始建立自己的決策支援系統；到 1990 年代，由於主從式架構、平行計算技術以及網際網路的興起，造就了資訊可被快速傳播、處理的時代，獲取資料已不成問題，如何運用及整合所得到的資料成為了重點，資料倉儲的概念因而產生(李卓翰，2003)。

資料倉儲是由 Bill Inmon 於 1980 年代晚期所提出，其認為資料倉儲是「支援管理決策過程、主題導向、經由整合、依時間而不同且不會劇變的資料集合。」(鄭旺昌，2012)，其為結合多個資料來源的中央儲存處，企業可利用資料倉儲技術儲存歷史資料，延伸更多訊息的資料報表，像是年度或季度的比較資訊，在商業智慧上是很重要的組成元素，可以整合、篩選與查詢企業的財務相關資料(傅梓瑜，2017)。

依據上述，資料倉儲具有主題導向、整合性、時間差異性與資料不變動性這四種特性，以下為這四種特色的介紹(黃熙勳，2013)：

### 1. 主題導向(Subject-Oriented)

將來源資料經由前處理過後，產生不同資訊，依據特定主題存放在資料庫內，以便於後續能隨時進行資料探勘。

### 2. 整合性(Integrated)

整合了不同平台且多種類的資料庫，藉由資料轉換的程序，將欄位名稱、變數、編碼方式、日期時間等屬性轉換成具有一致性的格式。

### 3. 時間差異性(Time-Variant)

為維持資料倉儲之可用性，須於特定時間點至各資訊系統中取得新資料，以確保資料倉儲資料之時效性。

### 4. 資料不變動性(Nonvolatile)

資料倉儲內的資料為具有歷史性之資料，具有不斷增加、不易變動等特性，以利於長時間性分析使用。



### 2.4.1. 星狀綱要結構與雪花綱要結構

資料倉儲主要分為事實資料表(Fact Table)與維度資料表(Dimension Table)建立關聯，每個維度資料表皆為獨立的維度，可藉由多個維度表組成多維模型，通常分為兩種架構，一個為星狀綱要結構，另一個為雪花綱要結構(黃熙勛，2013)。

星狀綱要結構是以事實資料表為中心，其欄位通常為主鍵與量值(Measure)，主鍵則分別與其他維度資料表具有關聯，如(圖 4)所示，而雪花綱要結構是星狀綱要的延伸，需透過正規化，是維度資料表間互相的關聯性，可減少資料重複與浪費空間(黃熙勛，2013)。

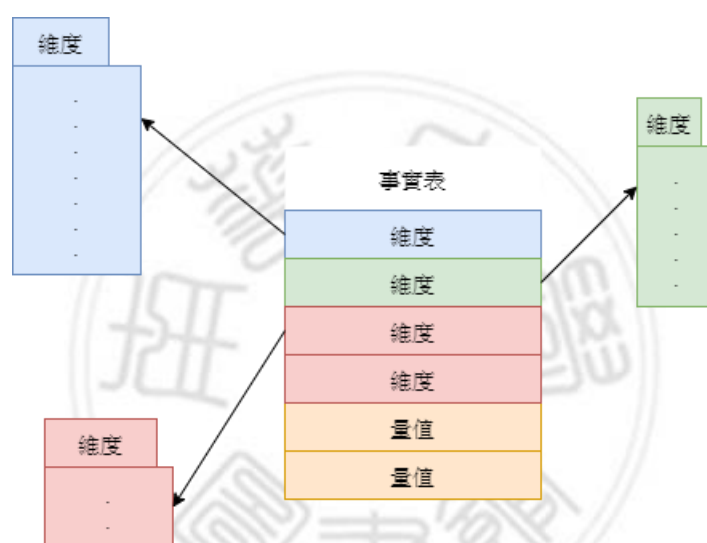


圖 4 星狀綱要結構圖

### 2.4.2. ETL(Extract-Transform-Load)

ETL(Extract-Transform-Load)是資料倉儲進行更新與建置時，對資料倉儲內資料執行萃取、轉換、載入之流程，ELT 通常使用在大量的資料或是複雜的邏輯轉換以及運算能力較強的資料庫(傅梓瑜，2017)。

1. Extract 為資料萃取，從不同資料來源處篩選出不同格式之資料數據。
2. Transform 為資料轉換，將萃取出之資料針對系統之需求進行資料轉換。
3. Load 為資料載入，將轉換完成之資料進行儲存，透過設計維度與量值，將維度表與事實表存到資料倉儲中。

### 2.4.3. 資料方體之切割應用

(圖 5)為某公司產品銷售資料方體圖，其可使用下鑽(drill-down)、上卷(roll-up)、切片(slice)、切塊(dice)與轉軸(pivot)等方式，瀏覽與使用資料方體內部資料，如以下三點(陳品辰，2009)：

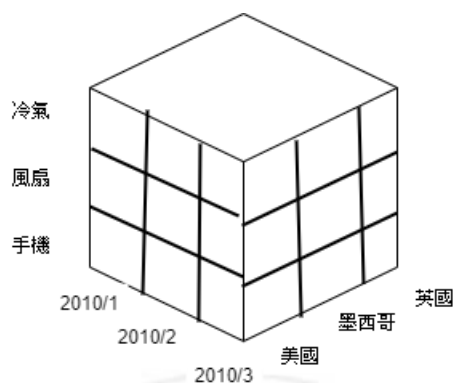


圖 5 某公司產品銷售資料方體圖

#### 1. 下鑽(drill-down)與上卷(roll-up)

資料使用者可往更細或更粗的層次進行瀏覽資料，如(圖 6)所示，時間維度為月份，可往下鑽至週與日，也可上卷至季與年份。

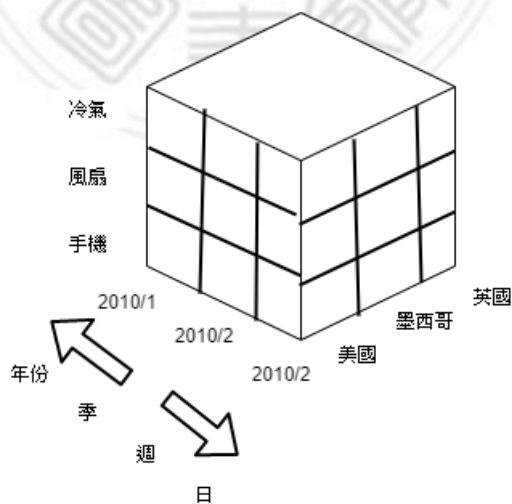


圖 6 某公司產品銷售資料方體之下鑽與上卷圖

## 2. 切片(slice)與切塊(dice)

切片為選擇兩個不同的維度資料進行切割與分析，如(圖 7)，而切塊為選擇三個以上不同的維度資料進行切割與分析，如(圖 8)。

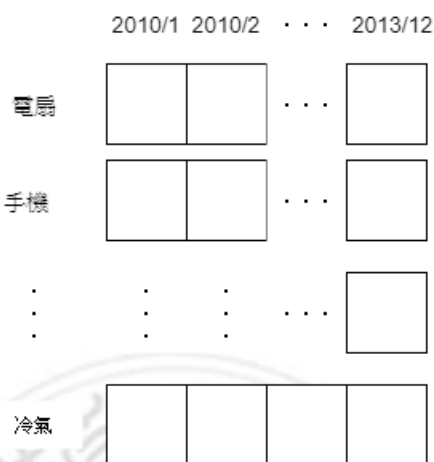


圖 7 某公司產品銷售資料方體切片圖

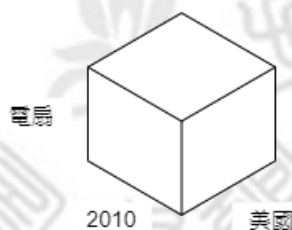


圖 8 某公司產品銷售資料方體切塊圖

## 3. 轉軸(pivot)

資料使用者可將維度進行軸的互換，以此搭配出不同的維度組合。

透過下鑽與上卷，資料使用者可向下與向上對不同維度層次的資料進行切割，取得多種不同維度層次的資料，而透過切片與切塊，資料使用者可選擇二維、三維甚至多維的切割方式，取得多種不同維度組合的資料。

## 第三章研究方法

首先建構本研究之研究流程，依照研究流程進行建置中醫門診資料庫與氣象因子資料庫，且對兩種資料以資料倉儲之萃取與轉換概念分別進行資料篩選與預處理，最後進行資料分群並處理極端數值，建構出具有資料倉儲概念之中醫氣象因子資料方體。

以下為本章所敘述的內容包含：

- 3.1. 研究流程
- 3.2. 中醫門診資料庫
- 3.3. 氣象因子資料庫
- 3.4. 資料篩選與預處理
- 3.5. 資料分群
- 3.6. 資料方體

### 3.1. 研究流程

本研究之研究架構依據研究動機目的與文獻探討之內容，將架構分成三大階段，如(圖 9)所示：

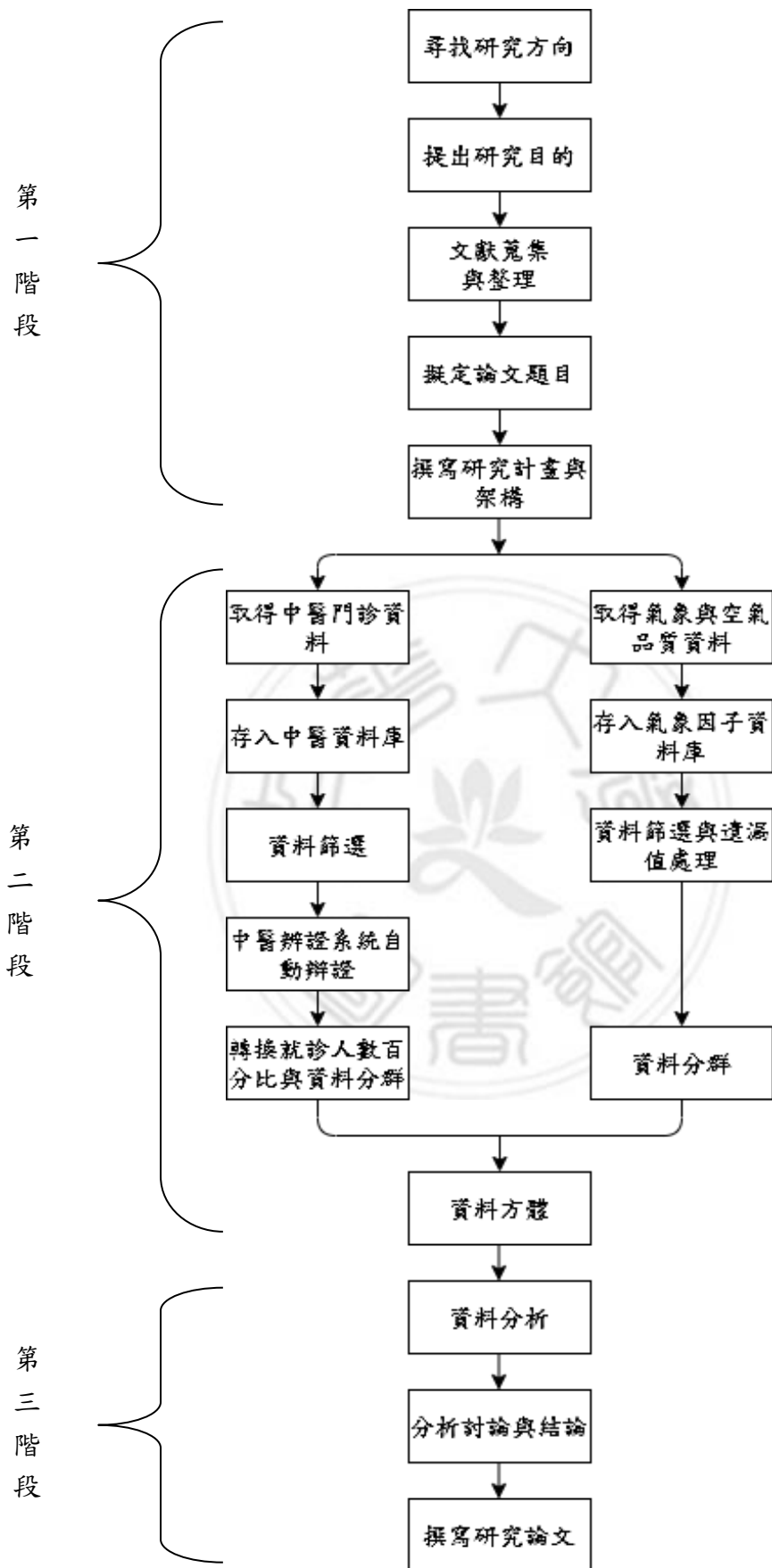


圖 9 研究流程圖

第一階段為題目定義期，藉由多方查閱論文與吸收網路新知，找尋研究方向，當找尋到合適且具有研究性質之事物時，整合所尋之研究方向且提出研究目的。而為了驗證研究目的之合理性，須針對特定之文獻進行蒐集與整理，最後將全部資料統整完成後，擬定確切之論文題目，進而撰寫完整之研究計畫與架構。

第二階段為原始資料蒐集期，本研究建構資料方體與資料探勘分析所需資料為癌症病患中醫辨證與氣象因子，故預計須分成兩部份進行資料蒐集。

第一部份為癌症病患中醫證候部份，首先取得嘉義大林慈濟醫院中醫腫瘤門診病歷原始資料，存入中醫門診資料庫，再篩選所需之中醫門診原始資料後，使用中正大學團隊所研發之中醫辨證系統，依照病患症狀轉換成證候，將每月證候之就診人數轉換所需之資料型態，進行分群後，存入資料方體中。

第二部份為氣象因子部份，由於本研究所使用之氣象因子包含中央氣象局與空氣品質網之資料，故先進行資料蒐集，取得氣象與空氣品質資料後，篩選所需之氣象與空氣品質資料，且將遺漏值處理完後，整合為氣象因子資料，最後將資料分群，存入資料方體中。

第三階段為實作完成期，對中醫證候與氣象因子資料方體進行不同維度之切割，先進行敘述性統計，再使用決策樹分析得出相關規則，分析結果與規則，進而得出結論，撰寫最後之研究論文。

### 3.2. 中醫門診資料庫

中醫原始門診資料是中正大學研究團隊從嘉義大林慈濟醫院中醫部取得中醫腫瘤門診病歷資料後，先行整理為原始資料之資料庫。

其中包含 2000 年至 2013 年之嘉義大林慈濟醫院中醫部每日就診之資料，資料表中包含就診日、住址、主治醫師、診斷 1~5 與治療藥品等等之病患相關資料，利用 Microsoft SQL Server 的匯入資料功能，將資料庫匯入中醫門診資料庫。

### 3.3. 氣象因子資料庫

本研究所使用之氣象因子資料包含氣象資料與空氣品質資料，故從中央氣象局與行政院環境保護署之環境資源資料庫取得資料，中央氣象局提供觀測資料查詢，讓大眾能夠方便且快速地取得相關氣象資料，且資料格式也分日、月與年三種模式，行政院環境保護署則提供環境資料庫供大眾使用，能取得日與月的空氣品質資料。

將氣象與空氣品質資料進行整合，利用 Microsoft SQL Server 的匯入資料功能，將 EXCEL 檔案匯入為氣象因子資料庫。

#### 3.3.1. 氣象資料

氣象資料可從中央氣象局之觀測資料查詢系統查詢，本研究所採用氣象資料為嘉義測站之每月資料，故測站設定為嘉義測站，資料格式為年報表才能取得一年 12 個月的資料，將 2010 年至 2013 年所有之資料下載完成，並整合成氣象資料 EXCEL 檔案。

#### 3.3.2. 空氣品質資料

空氣品質資料可從行政院環境保護署之環境資源資料庫查詢，本研究所採用空氣品質資料為嘉義地區之每月資料，將 2010 年至 2013 年所有之資料下載完成，並整合成空氣品質資料 EXCEL 檔案。

### 3.4. 資料篩選與預處理

基於上述操作，中醫門診資料庫與氣象因子資料庫內已有原始資料，分別針對兩個資料庫內之資料進行篩選與預處理，將資料型態轉換成本研究所須之型態，以利於後續儲存至資料倉儲內。

基於氣象因子影響人體速度並非即時性，其需要經由較長時間影響才較為容易出現症狀，且僅使用每日所累積之證候人數也較為不足，故本研究以月作為單位累積證候人數區間，並以累積證候人數區間前一個月做為氣象因子區間，如(圖 10)。



圖 10 氣象因子區間與累積證候人數區間

#### 3.4.1. 中醫門診資料篩選

中醫門診資料來自嘉義大林慈濟醫院中醫腫瘤門診，原始門診資料從 2000 年 1 月開始累積至 2013 年 12 月，本研究篩選出 2010 年 1 月至 2013 年 12 月且住址為雲嘉南地區的資料後，再將欄位進行整合，如(表 一)，總共有 38,608 筆資料。

表 一 中醫門診資料欄位篩選表

欄位	欄位目的
性別	病人基本資訊
年齡	
就醫日期	用於累計一段時間內的病歷數量
主述	病患所描述之症狀
客觀	經由醫師判斷之症狀
ICD-9 代碼	病患明確症狀代碼
住址	用來對應氣候資料的地理位置



### 3.4.2. 中醫門診資料預處理

經由上述資料篩選已完成資料倉儲的萃取部份，但本研究為建構具有資料倉儲概念之系統，篩選後之資料未經過轉換並不適合資料探勘使用，且原始中醫門診資料並未擁有證候欄位，故使用中正大學研究團隊所研發之中醫辨證系統，將資料轉換為合適之中醫證候。

中醫證候為眾多不同且具有相互關聯之症狀所組合而成，透過四診診斷病患之證候，針對證候對症下藥，前述之中醫門診資料庫所包含的主述，即為四診之結果。

中醫辨證系統是將與證候對應之原始症狀進行標準化後，將其取做標準症狀，整理出 20 種虛證與 47 種實證成為依據，再進行辨證，如(圖 11)，將此流程進行自動化，再根據模糊理論計算出每一種證候的歸屬度分數。且由於該系統僅接受中文，故會將英文替換成空字串，而標點符號會被當作是分詞的依據，替換成中醫辨證系統能夠認得的分詞符號。

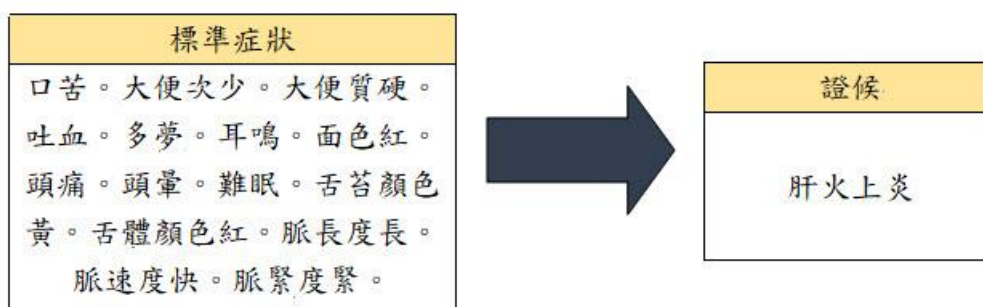


圖 11 中醫標準症狀對應證候

經由中醫辨證系統成功辨證之資料，共有 19065 筆病歷，88156 筆證候，由於中醫資料為每天就診之資料，為配合氣象因子之月報表資料格式，故將辨證後之資料以月份為單位進行整合，且為了供後續敘述性統計觀察，將資料依序處理為日、週、每年各月、特定月份、每年各季、特定季節、每年六氣、特定六氣與年份。

考量到就診人數持續上升，若單以每月人數進行資料探勘分析，會有誤差，故進行就診人數百分比轉換，將每個月證候人數除上該年證候總人數，得出該月患有此證候之人數百分比，以利後續資料探勘分析使用。

### 3.4.3. 氣象因子資料篩選

氣象因子資料庫內之資料分別為從中央氣象局與行政院環境保護署中所取得之原始資料，以嘉義地區 2010 年 1 月至 2013 年 12 月之氣候與空氣品質資料為主，依據中醫學說之外感六淫為基礎，針對 48 個氣象因子欄位進行篩選，篩選出 15 個適合之氣象因子欄位，如(表 二)，最後透過計算氣象區間之最高氣溫減去最低氣溫得出溫差，總欄位共 16 個。

表 二 氣象因子篩選表

六氣	判斷依據	篩選欄位
風	皮膚為空氣中致病因子進入身體之部位，故選擇會與皮膚接觸之欄位。	平均風速(m/s)
		測站平均氣壓(hPa)
		測站最高氣壓(hPa)
		測站最低氣壓(hPa)
		PM10( $\mu\text{g}/\text{m}^3$ )
		PM2.5( $\mu\text{g}/\text{m}^3$ )
寒	選擇與寒冷或炎熱特性有關之欄位。	平均氣溫( $^{\circ}\text{C}$ )
暑		最高氣溫( $^{\circ}\text{C}$ )
		最低氣溫( $^{\circ}\text{C}$ )
火		累積日照時數(hr)
濕	選擇與潮濕或乾燥特性有關之欄位。	平均相對溼度(%)
		累積降水量(mm)
燥		累積降水時數(hr)
		累積降水日數(天)
		最大日降水量(mm)

### 3.4.4. 氣象因子資料遺漏值處理

由於氣象資料與空氣品質資料是以感測器進行偵測，故有可能出現感測器失靈等等狀況，導致特定某些時段並無資料紀錄，故本研究在進行資料分群前，使用線性內插法先行補齊氣象因子資料之遺漏值，避免後續進行分析時出現誤差。

### 3.5. 資料分群

經由上述篩選與預處理，中醫資料與氣象因子資料已轉換成本研究所需之數值型態，為增加後續敘述性統計與資料探勘分析之多樣性，將已轉換好資料利用 RapidMiner 資料探勘軟體的 k-Means 分群法進行資料分群，轉變成類別型態，以利於後續建構資料方體、敘述性統計與資料探勘分析時，能擁有不同資料層次與切割方式。

首先將中醫資料與氣象因子資料匯入 RapidMiner 資料探勘軟體中，將 k-Means 分群法的 k 係數調整為 3，表示將資料分為三群，透過(圖 12)之流程，將中醫證候的實證與虛證共 67 種與 16 個氣象因子欄位依序進行 k-Means 分群法，如(圖 13)所示，將結果分為高、中、低三群，轉換成類別型態，其中 cluster\_1 為高、cluster\_0 為中、cluster\_2 為低。

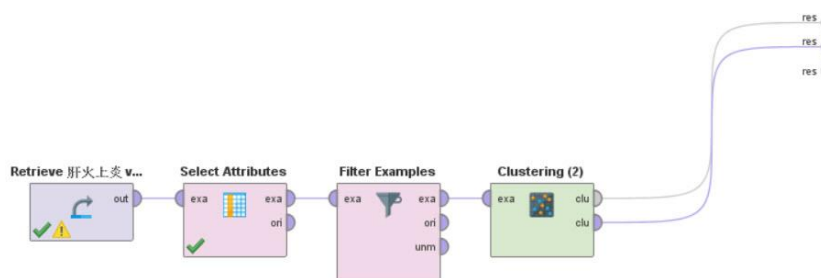


圖 12 RapidMiner 資料探勘軟體之 k-Means 分群法流程

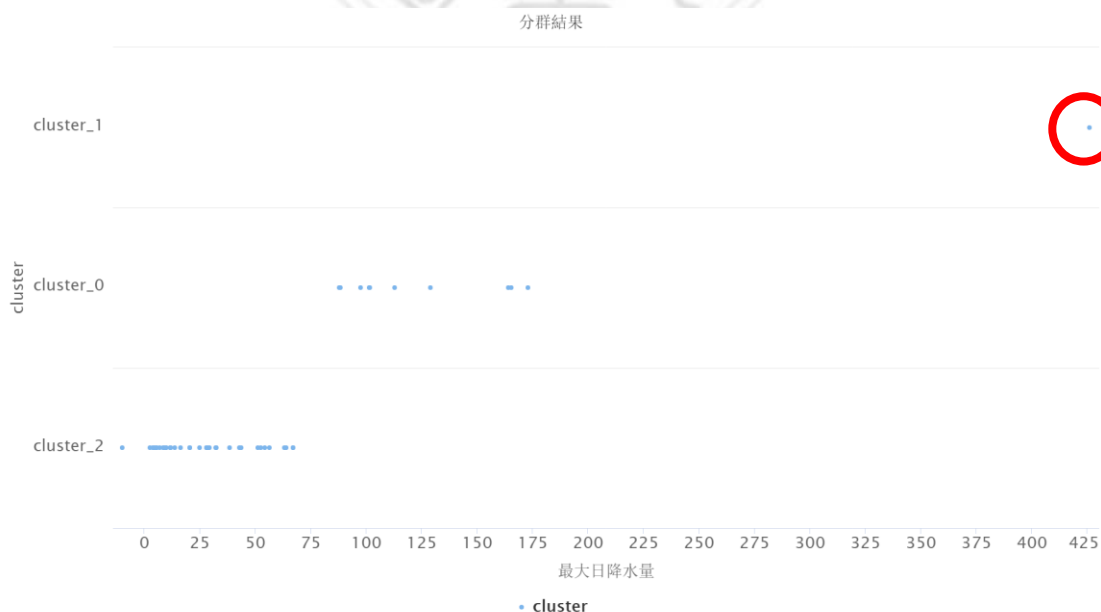


圖 13 RapidMiner 資料探勘軟體之 k-Means 分群法結果-以最大日降水量為例

但由(圖 13)之紅圈可發現，cluster\_1 僅有一個值，且此資料遠高於其他值，為極端數值，故須對其進行處理，故利用 RapidMiner 資料探勘軟體之 Filter Examples，將最大日降水量設為小於 200，將過高之極端數值先行排除，如(圖 14)所示。

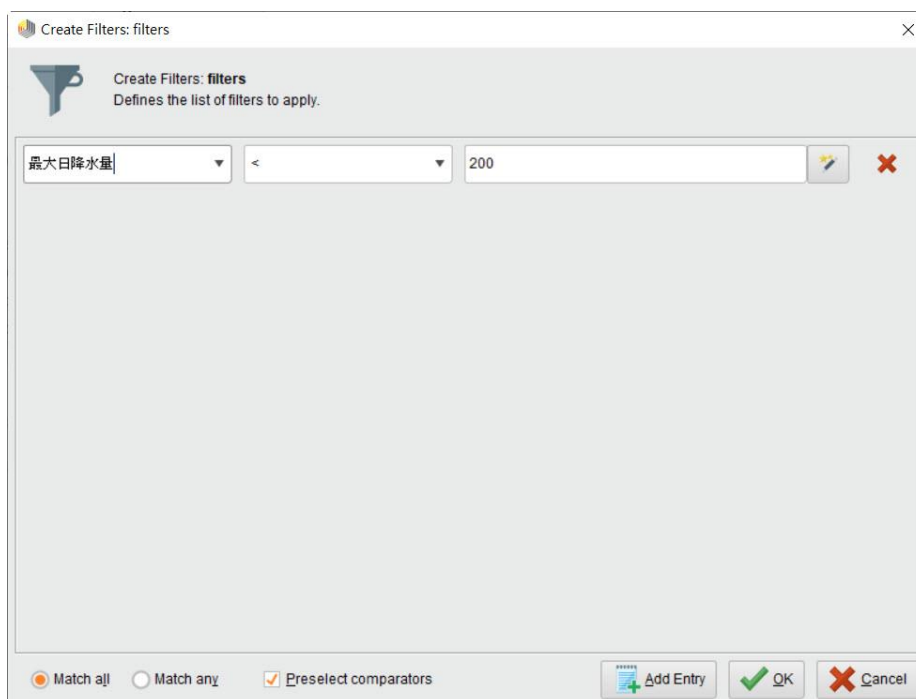


圖 14 RapidMiner 資料探勘軟體之 Filter Examples 應用圖

排除極端數值後，再次進行 k-Means 分群法，如(圖 15)之紅圈所示，cluster\_1 的數量雖上升不多，但皆在正常範圍內，最後將原先過高之極端數值列入類別高之分群中。

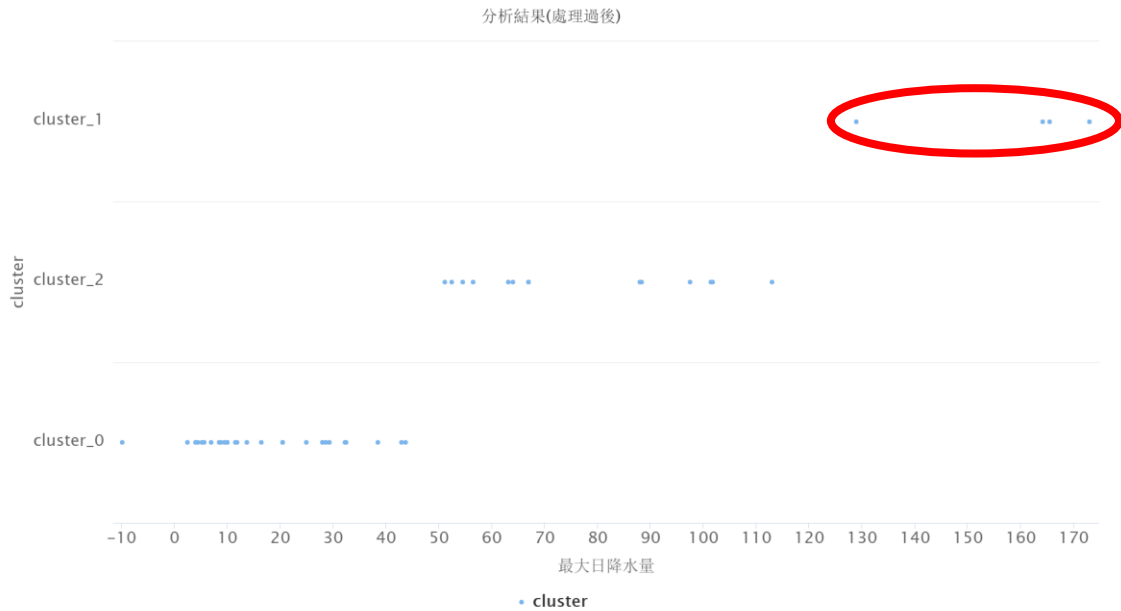


圖 15 RapidMiner 資料探勘軟體之 k-Means 分群法結果(處理過後)

經由上述極端值處理，處理前後範圍差異如(表 三)所示，處理過後範圍相較於處理前範圍，分群範圍有縮小，且群與群之間也變較為靠近，故此處理方式為合理的。

表 三 k-Means 分群法處理前後範圍差異表

分群	處理前範圍	處理後範圍
低	-9.8 ~ 67	-9.8 ~ 43.7
中	88 ~ 173	51.2 ~ 113
高	426	129 ~ 173、426

### 3.6. 資料方體

經由本章節之流程操作，中醫證候與氣象因子資料已經過資料篩選與預處理，成為適合進行資料探勘之資料，其資料格式如(表 四)，最終存入資料方體中。

表 四 資料探勘分析之資料

資料欄位	資料內容
日期(月份)	證候紀錄的日期
測站平均氣壓(hPa)	氣象因子區間內的氣壓平均值
測站最低氣壓(hPa)	氣象因子區間內的氣壓最高值
測站最高氣壓(hPa)	氣象因子區間內的氣壓最低值
平均氣溫(°C)	氣象因子區間內的氣溫平均值
最高氣溫(°C)	氣象因子區間內的氣溫最高值
最低氣溫(°C)	氣象因子區間內的氣溫最低值
溫差(°C)	氣象因子區間內的最高氣溫與最低氣溫之溫差值
平均風速(m/s)	氣象因子區間內的風速平均值
累積降水量(mm)	氣象因子區間內的每日平均降水量累計值
累積降水時數(hr)	氣象因子區間內的每日平均有降水的時數
累積降水日數(天)	氣象因子區間內的每月平均有降水的天數
最大日降水量(mm)	氣象因子區間內的每月最大日的降水量
平均相對溼度(%)	氣象因子區間內的相對溼度平均值
累計日照時數(hr)	氣象因子區間內的每日平均有陽光照射的時數
PM10( $\mu\text{g}/\text{m}^3$ )	氣象因子區間內的每日平均懸浮微粒 PM10
PM2.5( $\mu\text{g}/\text{m}^3$ )	氣象因子區間內的每日平均細顆粒懸浮為例 PM2.5
就診人數百分比(%)	累積證候人數區間內罹患證候的人數百分比
就診人數(人)	累積證候人數區間內罹患證候的人數

本研究透過資料倉儲之萃取、轉換與載入之步驟，完成以中醫證候與氣象因子為例之資料方體圖，如以下四點與(圖 16)所示：

1. 蒐集三種不同來源之中醫門診、氣象與空氣品質資料並儲存至資料庫
2. 資料篩選為從不同來源之資料庫篩選不同資料格式之資料，可視為萃取部份。
3. 中醫辨證、遺漏值處理與資料分群則依據分析系統之需求進行轉換資料格式，可從不同角度對資料方體進行應用，可視為轉換部份。
4. 將轉換完成之資料設計維度與量值並完成事實資料表與維度資料表，最終存入中醫證候與氣象因子資料方體，可視為載入部份

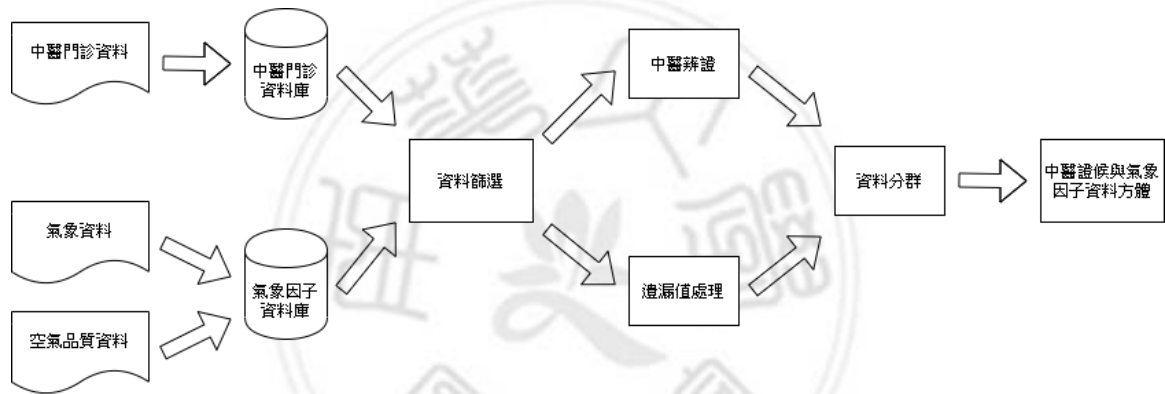


圖 16 資料方體圖-以中醫證候與氣象因子為例

由上述可知，本研究整合三種不同來源資料，並建構出資料方體，利用資料方體對於顆粒的粗細變化，進行不同維度的切割，為後續資料探勘分析建立便利地取得不同維度資料之管道。

資料方體是由維度與量值組合而成的多維度資料模型，依照不同維度建立起來，維度包含屬性與階層，而中醫證候與氣象因子資料方體是依據資料倉儲之萃取、轉換與載入之概念，其星狀綱要結構包含中醫證候維度、時間維度與各種氣象因子維度，而各維度中又有不同之階層，如(圖 17)所示。

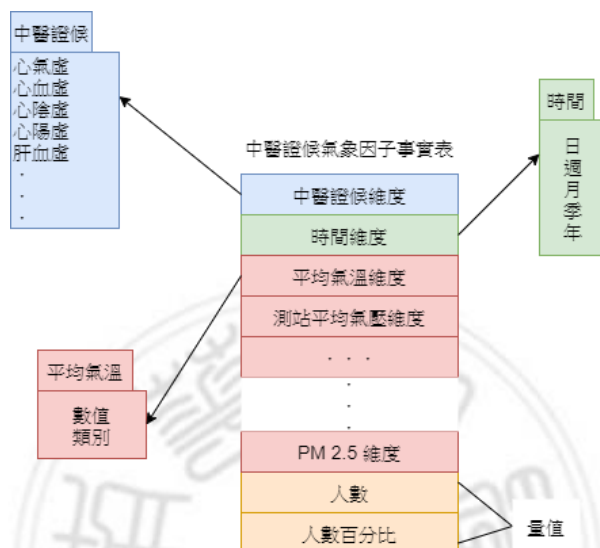


圖 17 中醫證候與氣象因子資料方體之星狀綱要結構



## 第四章資料分析

經由第三章後，本研究已建構出具有資料倉儲概念之資料方體，依據陳宗仁(2019)的研究結果，對資料方體進行不同的處理方式。

首先針對全部證候進行敘述性統計，而肝火上炎、心脈痺阻(氣滯)、心脈痺阻(寒凝)與肝陽上亢做為主要目標，藉由切割不同的資料維度，進行資料視覺化觀察，探討在不同的維度切割能找出何種規律，並挑選出除主要目標以外，表現較佳之證候一同呈現與討論。

接著以肝火上炎證候為主要目標，再次切割不同的資料維度取出所需之資料，使用RapidMiner 進行決策樹分析，透過不同準則、不同訓練資料集比例與是否剪枝找出合適不同輸入變數之模型，並將所得出之規則以陳宗仁(2019)的研究結果進行比較、結果討論與分析。

#### 4.1. 敘述性統計

本節針對中醫證候與氣象因子資料方體進行六種不同的維度切割，分別為：中醫證候維度與時間維度(每年各月)切割、中醫證候維度與時間維度(特定月份)切割、中醫證候維度與時間維度(每年各季)切割、中醫證候維度與時間維度(特定季節)切割、中醫證候維度與時間維度(每年六氣)切割與中醫證候維度與時間維度(特定六氣)切割。

在中醫證候維度切割資料時，以 2 種不同證候與陳宗仁研究結果中具有關係的 4 種證候為主要觀察對象，而進行敘述性統計時，發現有 16 種證候其就診人數為 0 人，有 2 種證候則就診人數大部份為 0 人，最後有 6 種證候則就診人數幾乎低於 10 人，合計共 24 種證候並不進行敘述性統計。

##### 4.1.1. 中醫證候維度與時間維度(每年各月)切割

首先從中醫證候與氣象因子資料方體內切割所需資料，如(圖 18)與(圖 19-26)所示，以資料方體切片之方式依序切出中醫證候之各個階層與時間維度之每年各月階層，透過觀察中醫證候與每年各月之資料，找出中醫證候在每年不同月份的關係。

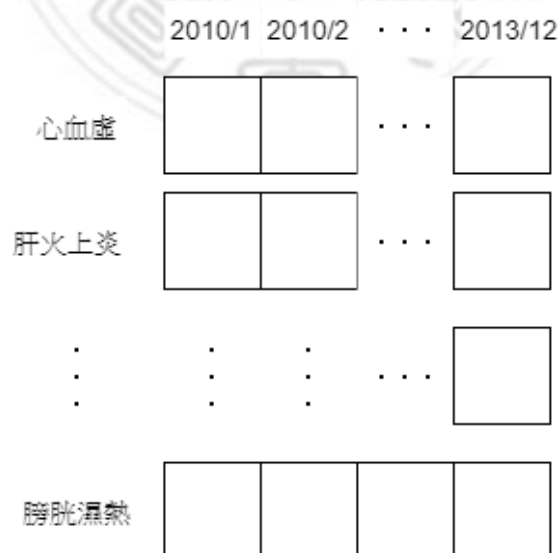


圖 18 中醫證候維度與時間維度(每年各月)之資料方體切片

觀測時間	2010/1	2010/2	2010/3	2010/4	2010/5	2010/6	2010/7	2010/8	2010/9	2010/10	2010/11	2010/12
心氣虛	20	25	33	32	20	24	34	31	26	22	30	41
心血虛	25	20	27	21	24	29	35	45	34	51	48	45
心陰虛	24	17	31	33	35	39	32	46	49	66	49	50
心陽虛	1	1	0	3	0	2	2	4	4	7	2	8
肝血虛	3	5	4	4	7	4	9	5	4	11	14	20
脾氣虛	29	37	39	69	38	45	49	53	62	65	65	87
脾陽虛	2	1	0	4	4	6	3	5	6	8	5	7
脾氣陷	7	9	10	11	7	10	9	9	7	21	17	19
胃陰虛	28	16	31	49	40	38	30	40	50	60	65	70
肺氣虛	25	28	29	38	23	24	30	31	36	29	32	49
肺陰虛	6	2	5	8	6	12	5	10	11	19	14	15
腎精不足	7	4	9	12	5	4	7	7	9	11	9	8
腎不納氣	19	21	23	29	20	25	24	24	23	21	21	22
腎氣不固	32	41	40	51	45	49	64	71	66	68	82	88
腎陰虛	17	9	14	26	29	32	26	35	41	50	46	50
腎陽虛	22	22	26	28	20	30	44	50	42	58	56	54
大腸陽虛	8	5	7	11	5	15	9	16	11	10	20	19

圖 19 中醫證候(虛證)與時間維度(2010 年各月)之資料切片結果

觀測時間	2010/1	2010/2	2010/3	2010/4	2010/5	2010/6	2010/7	2010/8	2010/9	2010/10	2010/11	2010/12
寒襲衛表	8	8	10	16	17	11	10	12	12	14	16	23
暑襲衛表	18	6	17	18	14	14	14	18	16	26	31	41
燥襲衛表	15	4	8	12	16	10	9	9	14	23	21	39
熱襲衛表	15	4	7	14	12	15	16	18	17	19	24	42
脈痺阻(寒)	12	8	7	12	12	10	8	18	16	25	17	18
心火亢盛	7	4	14	20	14	15	7	11	11	25	27	21
脈痺阻(痰)	18	19	28	25	19	24	36	45	30	30	39	46
火擾心(鬱)	19	11	22	22	26	26	23	25	32	55	40	60
火擾心(躁)	10	6	11	4	15	10	14	9	11	9	10	7
擾心(外感)	13	17	35	35	38	37	28	38	47	52	50	42
脈痺阻(血)	0	0	1	4	2	5	4	3	3	3	0	4
脈痺阻(氣)	13	11	17	14	12	17	17	24	13	7	11	9
寒邪客肺	6	1	3	4	3	3	4	6	7	14	6	9
痰濕阻肺	9	6	8	7	10	10	12	20	16	13	14	21
痰熱壅肺	10	7	7	8	5	7	6	11	11	11	7	18
大腸液虧	19	19	34	31	35	32	30	22	24	40	44	30
大腸實熱	8	6	11	14	11	13	12	17	19	25	30	27
寒邪客脾	4	2	4	3	3	2	4	2	1	6	3	12
熱邪客脾	10	11	10	20	10	10	10	16	24	13	20	33
胃氣上逆	12	8	8	18	20	9	14	19	22	15	28	40
食滯胃脘	6	4	6	5	12	4	10	7	9	7	21	12
寒滯肝脈	10	8	5	11	9	10	4	13	13	18	18	18
肝火上炎	18	17	33	40	42	34	39	38	43	52	60	51
肝陽上亢	18	14	21	33	23	36	34	51	56	63	62	74
肝氣鬱結	0	1	0	0	3	1	1	1	1	2	3	1
膽郁痰擾	19	11	21	27	18	19	20	43	32	45	38	48

圖 20 中醫證候(實證)與時間維度(2010 年各月)之資料切片結果

觀測時間	2011/1	2011/2	2011/3	2011/4	2011/5	2011/6	2011/7	2011/8	2011/9	2011/10	2011/11	2011/12
心氣虛	33	27	49	51	71	53	56	50	71	58	65	66
心血虛	36	26	54	69	60	64	74	65	76	69	87	72
心陰虛	48	35	42	67	57	65	67	59	55	66	94	68
心陽虛	0	2	3	13	13	11	16	16	20	12	12	9
肝血虛	16	13	14	13	10	6	9	18	15	19	15	7
脾氣虛	59	44	66	69	92	83	93	95	115	120	152	133
脾陽虛	0	4	4	5	6	9	20	3	15	8	9	7
脾氣陷	12	16	26	29	36	28	25	17	26	18	36	25
胃陰虛	49	43	60	69	49	53	56	73	54	58	65	57
肺氣虛	33	28	48	47	54	55	57	55	55	52	62	72
肺陰虛	12	8	18	24	16	19	13	17	12	10	9	11
腎精不足	5	7	9	12	5	4	6	4	6	8	9	3
腎不納氣	20	18	36	29	35	32	42	30	55	48	56	41
腎氣不固	72	42	70	50	72	70	63	51	75	62	85	82
腎陰虛	35	30	34	62	48	62	58	62	43	48	63	47
腎陽虛	48	38	56	70	63	55	58	62	57	52	90	73
大腸陽虛	17	17	26	24	31	12	31	21	19	25	47	37

圖 21 中醫證候(虛證)與時間維度(2011 年各月)之資料切片結果

觀測時間	2011/1	2011/2	2011/3	2011/4	2011/5	2011/6	2011/7	2011/8	2011/9	2011/10	2011/11	2011/12
寒襲衛表	23	12	19	36	38	22	25	24	26	29	44	44
暑襲衛表	32	22	20	31	32	30	25	31	30	31	45	49
燥襲衛表	22	28	26	36	31	22	20	23	23	18	27	34
熱襲衛表	30	20	34	31	21	16	16	34	27	29	38	31
脈痺阻(寒)	19	13	20	47	43	45	50	46	56	37	58	52
心火亢盛	21	13	24	23	21	29	31	23	29	30	46	26
脈痺阻(痰)	47	31	39	64	54	66	73	55	77	63	86	86
火擾心(鬱)	45	35	31	40	33	19	35	46	37	38	52	35
火擾心(躁)	11	9	7	6	7	7	5	5	4	6	7	7
擾心(外感)	42	30	40	44	42	63	55	55	56	44	56	47
脈痺阻(血)	1	9	7	5	15	8	10	13	16	8	13	18
脈痺阻(氣)	10	9	20	35	38	41	40	43	49	31	34	35
寒邪客肺	4	6	10	17	12	8	4	8	8	5	15	22
痰濕阻肺	12	12	25	21	24	10	13	21	22	26	24	28
痰熱壅肺	9	14	12	22	15	6	13	18	11	16	19	20
大腸液虧	22	23	48	69	49	54	48	61	59	37	54	37
大腸實熱	19	20	23	22	23	28	33	27	33	34	39	29
寒邪客脾	2	5	9	5	5	7	7	8	4	6	6	7
熱邪客脾	19	14	25	27	32	19	37	29	7	24	34	28
胃氣上逆	17	23	33	25	26	14	22	35	23	26	28	32
食滯胃脘	6	4	13	17	12	15	9	10	9	10	13	11
寒滯肝脈	15	14	14	35	23	20	30	17	30	20	35	31
肝火上炎	48	39	57	71	63	66	78	73	85	70	82	81
肝陽上亢	56	41	49	79	54	63	68	74	74	59	66	72
肝氣鬱結	1	6	6	8	10	9	13	9	17	13	11	6
膽郁痰擾	41	33	60	62	50	50	40	44	44	52	71	46

圖 22 中醫證候(實證)與時間維度(2011 年各月)之資料切片結果

觀測時間	2012/1	2012/2	2012/3	2012/4	2012/5	2012/6	2012/7	2012/8	2012/9	2012/10	2012/11	2012/12
心氣虛	41	63	35	49	61	63	69	53	46	65	85	77
心血虛	69	88	80	71	118	108	105	103	95	104	88	105
心陰虛	46	59	54	51	101	88	98	84	86	106	82	90
心陽虛	11	2	8	9	15	15	21	18	10	18	18	18
肝血虛	18	19	19	26	36	30	21	24	17	26	20	33
脾氣虛	92	117	107	114	164	147	128	107	107	156	123	118
脾陽虛	10	17	15	16	14	25	14	8	4	4	6	13
脾氣陷	22	47	43	26	33	36	45	38	29	34	49	42
胃陰虛	41	65	53	61	83	56	82	71	69	93	105	108
肺氣虛	45	67	50	43	56	51	58	39	58	68	64	55
肺陰虛	12	9	9	12	10	11	23	20	10	16	24	16
腎精不足	2	2	5	7	14	8	14	13	13	18	5	11
腎不納氣	30	38	31	36	50	55	52	39	39	55	50	45
腎氣不固	59	56	40	57	79	71	69	60	69	90	77	66
腎陰虛	34	43	41	47	73	68	90	66	90	102	86	81
腎陽虛	40	62	60	63	79	92	97	68	69	82	78	78
大腸陽虛	22	32	27	38	28	35	33	33	42	28	28	27

圖 23 中醫證候(虛證)與時間維度(2012年各月)之資料切片結果

觀測時間	2012/1	2012/2	2012/3	2012/4	2012/5	2012/6	2012/7	2012/8	2012/9	2012/10	2012/11	2012/12
寒襲衛表	20	35	26	38	25	23	43	36	39	46	48	37
暑襲衛表	34	43	42	45	37	23	38	31	40	54	65	47
燥襲衛表	24	34	35	37	27	23	30	20	34	68	55	29
熱襲衛表	24	26	29	35	32	32	44	40	37	48	49	43
脈痺阻(寒)	34	27	35	44	54	39	47	25	27	39	40	50
心火亢盛	24	44	42	29	35	48	43	40	51	56	45	48
脈痺阻(痰)	49	69	66	68	82	84	107	63	82	80	93	100
火擾心(鬱)	53	61	55	48	57	62	64	50	46	56	83	67
火擾心(躁)	10	21	10	17	14	15	12	14	9	19	30	20
擾心(外感)	43	73	66	64	84	66	63	73	82	98	96	80
脈痺阻(血)	4	6	9	12	18	6	11	8	11	7	4	16
脈痺阻(氣)	22	27	24	33	39	44	44	36	28	33	40	39
寒邪客肺	12	21	24	17	10	19	11	3	10	11	8	7
痰濕阻肺	19	34	48	31	27	27	30	10	39	35	22	30
痰熱壅肺	14	19	20	21	21	21	17	3	23	23	11	18
大腸液虧	34	49	42	42	73	53	73	54	68	92	95	89
大腸實熱	27	46	41	36	36	51	43	49	57	64	58	54
寒邪客脾	9	13	12	7	10	12	10	1	3	10	8	9
熱邪客脾	21	30	42	40	64	41	59	63	36	42	27	28
胃氣上逆	22	26	45	40	55	54	43	29	31	37	45	36
食滯胃脘	6	18	20	11	23	11	11	20	16	15	19	10
寒滯肝脈	25	21	25	29	29	20	29	18	20	28	21	32
肝火上炎	64	80	81	81	100	97	121	119	125	134	118	125
肝陽上亢	70	73	66	74	95	92	116	124	104	89	110	126
肝氣鬱結	5	7	7	9	7	6	8	6	4	6	15	11
膽郁痰擾	55	74	73	62	100	98	79	63	68	109	102	99

圖 24 中醫證候(實證)與時間維度(2012年各月)之資料切片結果

觀測時間	2013/1	2013/2	2013/3	2013/4	2013/5	2013/6	2013/7	2013/8	2013/9	2013/10	2013/11	2013/12
心氣虛	58	71	119	129	118	114	132	140	130	130	138	137
心血虛	115	87	132	153	135	120	160	161	163	183	192	183
心陰虛	95	73	86	122	113	85	133	112	106	130	143	121
心陽虛	17	15	20	35	22	24	16	16	14	22	30	17
肝血虛	47	23	36	40	35	33	48	39	41	40	41	41
脾氣虛	116	100	150	174	166	155	172	151	164	137	170	175
脾陽虛	11	14	20	22	13	18	23	18	35	33	43	35
脾氣陷	43	39	46	53	47	62	62	69	74	69	86	75
胃陰虛	111	83	119	120	102	89	120	132	93	123	120	110
肺氣虛	56	69	132	137	91	88	114	96	84	98	98	95
肺陰虛	23	12	35	38	41	41	39	62	38	51	63	50
腎精不足	10	10	8	17	26	18	23	23	26	25	23	30
腎不納氣	40	39	85	105	69	76	77	70	68	63	78	56
腎氣不固	61	70	127	130	118	107	100	109	108	90	101	89
腎陰虛	84	53	86	120	112	87	129	106	115	94	113	95
腎陽虛	82	63	105	121	126	98	116	114	107	105	114	90
大腸陽虛	41	26	48	41	45	30	40	54	57	68	79	78

圖 25 中醫證候(虛證)與時間維度(2013 年各月)之資料切片結果

觀測時間	2013/1	2013/2	2013/3	2013/4	2013/5	2013/6	2013/7	2013/8	2013/9	2013/10	2013/11	2013/12
寒襲衛表	44	38	52	51	48	45	51	53	59	56	80	72
暑襲衛表	39	36	68	101	82	58	59	46	54	33	46	39
燥襲衛表	52	29	55	70	72	47	56	47	49	43	39	42
熱襲衛表	38	43	52	87	79	44	66	36	33	47	53	65
脈痺阻(寒)	57	41	56	54	70	74	61	82	59	91	106	101
心火亢盛	61	36	46	43	46	42	59	51	40	46	49	60
脈痺阻(痰)	96	64	103	110	134	90	111	110	110	133	157	146
火擾心(鬱)	65	34	66	54	63	44	78	59	68	67	92	78
火擾心(躁)	29	10	17	13	12	4	34	29	15	22	32	30
火擾心(外感)	123	78	97	103	106	110	122	106	86	72	109	99
脈痺阻(血)	23	12	16	21	40	37	27	28	32	39	59	58
脈痺阻(氣)	43	36	46	60	81	64	71	93	78	94	98	84
寒邪客肺	13	9	18	19	14	10	17	16	20	9	4	11
痰濕阻肺	25	19	32	40	65	35	62	51	40	23	59	46
痰熱壅肺	14	15	22	32	45	21	35	31	28	16	40	28
大腸液虧	96	61	82	86	66	74	96	94	93	99	110	102
大腸實熱	77	47	57	41	38	47	62	49	41	37	48	61
寒邪客脾	9	7	11	16	20	17	22	37	27	28	24	30
熱邪客脾	39	33	46	42	51	52	56	53	36	56	36	54
胃氣上逆	37	41	67	54	73	72	58	60	43	54	62	37
食滯胃脘	20	18	30	16	13	18	33	27	17	16	27	26
寒滯肝脈	32	21	32	35	31	34	27	31	18	24	28	35
肝火上炎	141	88	128	119	124	110	148	165	146	163	163	172
肝陽上亢	108	73	116	131	130	111	125	145	149	169	180	165
肝氣鬱結	5	3	11	7	22	19	14	15	9	10	21	12
膽郁痰擾	89	62	101	84	99	96	119	127	107	103	129	111

圖 26 中醫證候(實證)與時間維度(2013 年各月)之資料切片結果

(圖 27)為心陰虛每年各月就診人數百分比折線圖，從圖可看出，心陰虛的就診人數百分比在每年 1-2 月皆偏低，且 2010、2012 年 10 月與 2011、2013 年 11 月偏高。

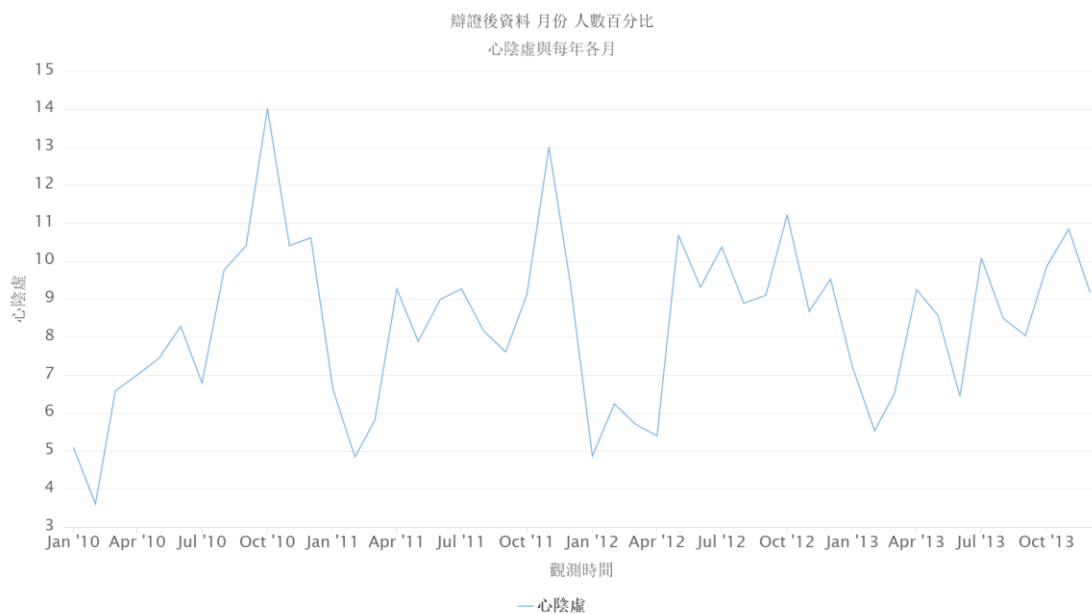


圖 27 心陰虛每年各月就診人數百分比折線圖

(圖 28)為腎不納氣每年各月就診人數百分比折線圖，從圖可看出，腎不納氣的就診人數百分比在 2011、2013 年 1-2 月偏低，且 2010、2013 年 4 月偏高。



圖 28 腎不納氣每年各月就診人數百分比折線圖

(圖 29)為心脈痺阻(寒凝)每年各月就診人數百分比折線圖，從圖可看出，心脈痺阻(寒凝)的就診人數百分比在每年1-2月皆偏低，且在2010年10月偏高。

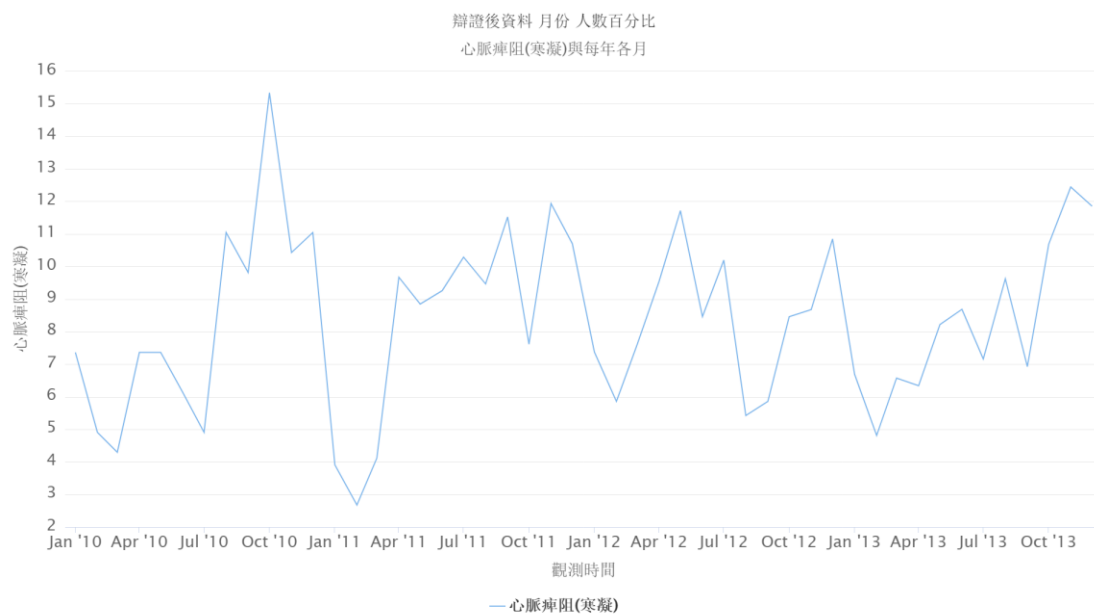


圖 29 心脈痺阻(寒凝)每年各月就診人數百分比折線圖

(圖 30)為心脈痺阻(氣滯)每年各月就診人數百分比折線圖，從圖可看出，心脈痺阻(氣滯)的就診人數百分比除2012年外，其餘年8-9月皆偏高。



圖 30 心脈痺阻(氣滯)每年各月就診人數百分比折線圖



(圖 31)為肝火上炎每年各月就診人數百分比折線圖，從圖可看出，肝火上炎的就診人數百分比在每年1-2月皆偏低，且在每年11-12月皆偏高。

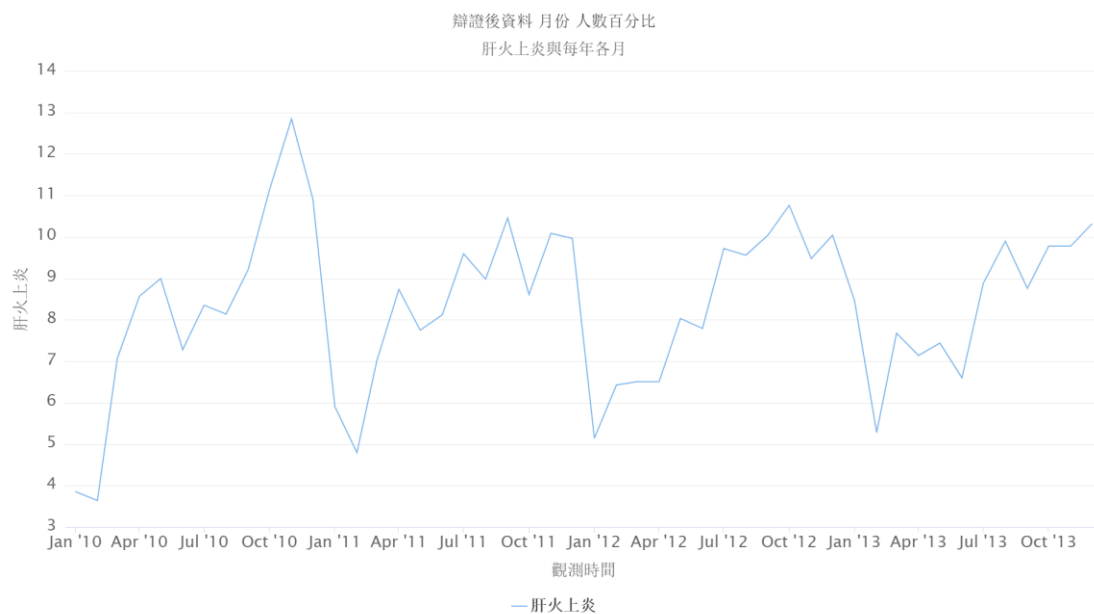


圖 31 肝火上炎每年各月就診人數百分比折線圖

(圖 32)為肝陽上亢每年各月就診人數百分比折線圖，從圖可看出，肝陽上亢的就診人數百分比在每年1-2月皆偏低，且在每年11-12月皆偏高。



圖 32 肝陽上亢每年各月就診人數百分比折線圖

#### 4.1.2. 中醫證候維度與時間維度(特定月份)切割

首先從中醫證候與氣象因子資料方體內切割所需資料，如(圖 33)與(圖 34-35)所示，以資料方體切片之方式依序切出中醫證候之各個階層與時間維度之特定月份階層，特定月份為將 2010-2013 年之各月就診人數相加起來除以 2010-2013 年之就診人數，即能取得 2010-2013 年各月份之就診人數百分比，透過觀察中醫證候與特定月份之資料，找出中醫證候在每個特定月份不同的關係。

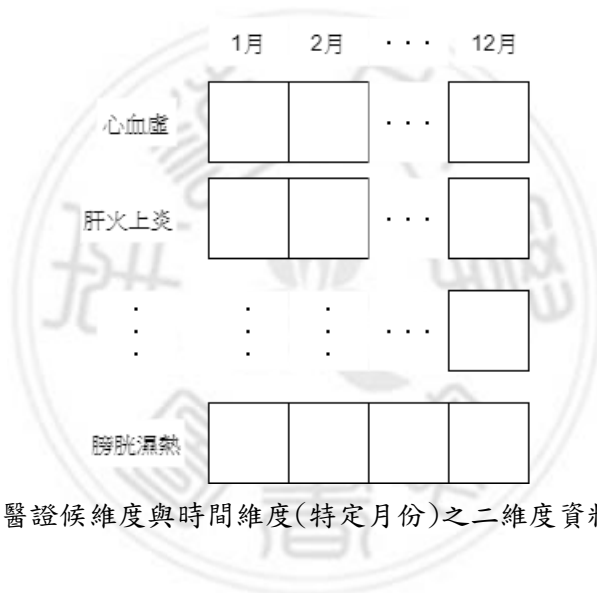


圖 33 中醫證候維度與時間維度(特定月份)之二維度資料方體切割

觀測時間	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
心氣虛	152	186	236	261	270	254	291	274	273	275	318	321
心血虛	245	216	293	314	337	321	374	374	368	407	415	405
心陰虛	213	184	213	273	306	277	330	301	296	368	368	329
心陽虛	29	20	31	60	50	52	55	54	48	59	62	52
肝血虛	84	60	73	83	88	73	87	86	77	96	90	101
脾氣虛	296	298	362	426	460	430	442	406	448	478	510	513
脾陽虛	23	36	39	47	37	58	60	34	60	53	63	62
脾氣陷	84	111	125	119	123	136	141	133	136	142	188	161
胃陰虛	229	207	263	299	274	236	288	316	266	334	355	345
肺氣虛	159	192	259	265	224	218	259	221	233	247	256	271
肺陰虛	53	31	67	82	73	83	80	109	71	96	110	92
腎精不足	24	23	31	48	50	34	50	47	54	62	46	52
腎不納氣	109	116	175	199	174	188	195	163	185	187	205	164
腎氣不固	224	209	277	288	314	297	296	291	318	310	345	325
腎陰虛	170	135	175	255	262	249	303	269	289	294	308	273
腎陽虛	192	185	247	282	288	275	315	294	275	297	338	295
大腸陽虛	88	80	108	114	109	92	113	124	129	131	174	161

圖 34 中醫證候(虛證)與時間維度(特定月份)之資料切片結果

觀測時間	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
寒襲衛表	95	93	107	141	128	101	129	125	136	145	188	176
暑襲衛表	123	107	147	195	165	125	136	126	140	144	187	176
燥襲衛表	113	95	124	155	146	102	115	99	120	152	142	144
熱襲衛表	107	93	122	167	144	107	142	128	114	143	164	181
心脈痺阻(寒凝)	122	89	118	157	179	168	166	171	158	192	221	221
心火亢盛	113	97	126	115	116	134	140	125	131	157	167	155
心脈痺阻(痰聚)	210	183	236	267	289	264	327	273	299	306	375	378
痰火擾心(鬱)	182	141	174	164	179	151	200	180	183	216	267	240
痰火擾心(躁)	60	46	45	40	48	36	65	57	39	56	79	64
痰火擾心(外感熱)	221	198	238	246	270	276	268	272	271	266	311	268
心脈痺阻(血瘀)	28	27	33	42	75	56	52	52	62	57	76	96
心脈痺阻(氣滯)	88	83	107	142	170	166	172	196	168	165	183	167
寒邪客肺	35	37	55	57	39	40	36	33	45	39	33	49
痰濕阻肺	65	71	113	99	126	82	117	102	117	97	119	125
痰熱壅肺	47	55	61	83	86	55	71	63	73	66	77	84
大腸液虧	171	152	206	228	223	213	247	231	244	268	303	258
大腸實熱	131	119	132	113	108	139	150	142	150	160	175	171
寒邪客脾	24	27	36	31	38	38	43	48	35	50	41	58
熱邪客脾	89	88	123	129	157	122	162	161	103	135	117	143
胃氣上逆	88	98	153	137	174	149	137	143	119	132	163	145
食滯胃脘	38	44	69	49	60	48	63	64	51	48	80	59
寒滯肝脈	82	64	76	110	92	84	90	79	81	90	102	116
肝火上炎	271	224	299	311	329	307	386	395	399	419	423	429
肝陽上亢	252	201	252	317	302	302	343	394	383	380	418	437
肝氣鬱結	11	17	24	24	42	35	36	31	31	31	50	30
膽郁痰擾	204	180	255	235	267	263	258	277	251	309	340	304

圖 35 中醫證候(實證)與時間維度(特定月份)之資料切片結果

(圖 36)為心陽虛特定月份就診人數百分比長條圖，從圖可看出，心陽虛的就診人數百分比在 2 月偏低，而在 11 月偏高。

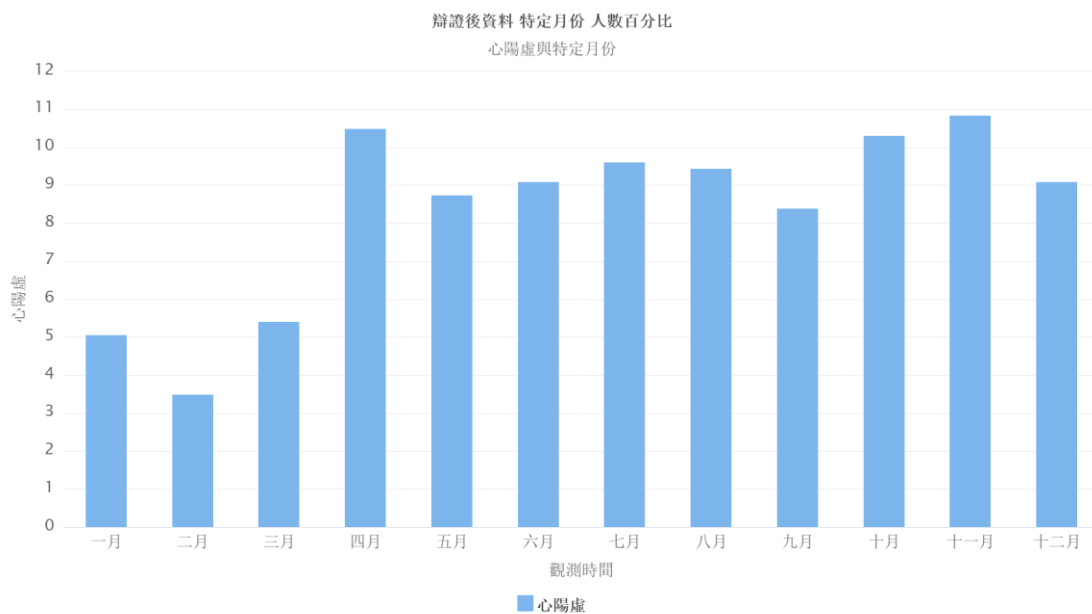


圖 36 心陽虛特定月份就診人數百分比長條圖

(圖 37)為腎不納氣特定月份就診人數百分比長條圖，從圖可看出，腎不納氣的就診人數百分比在 1 月偏低，而在 11 月偏高。

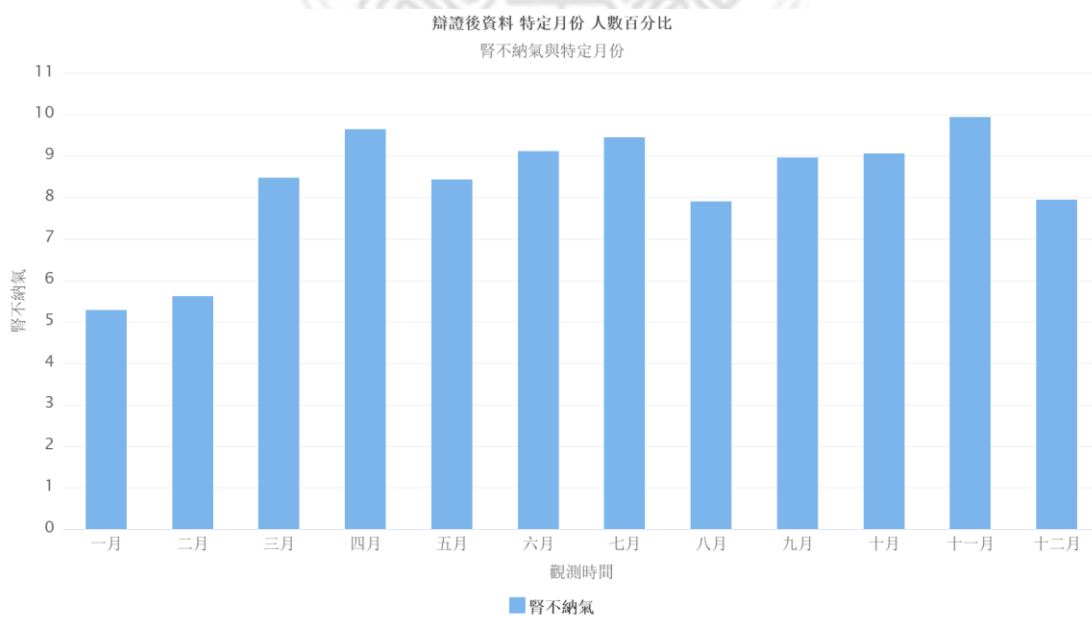


圖 37 腎不納氣特定月份就診人數百分比長條圖

(圖 38)為心脈痺阻(寒凝)特定月份就診人數百分比長條圖，從圖可看出，心脈痺阻(寒凝)的就診人數百分比在 1 月偏低，而在 11 月偏高。

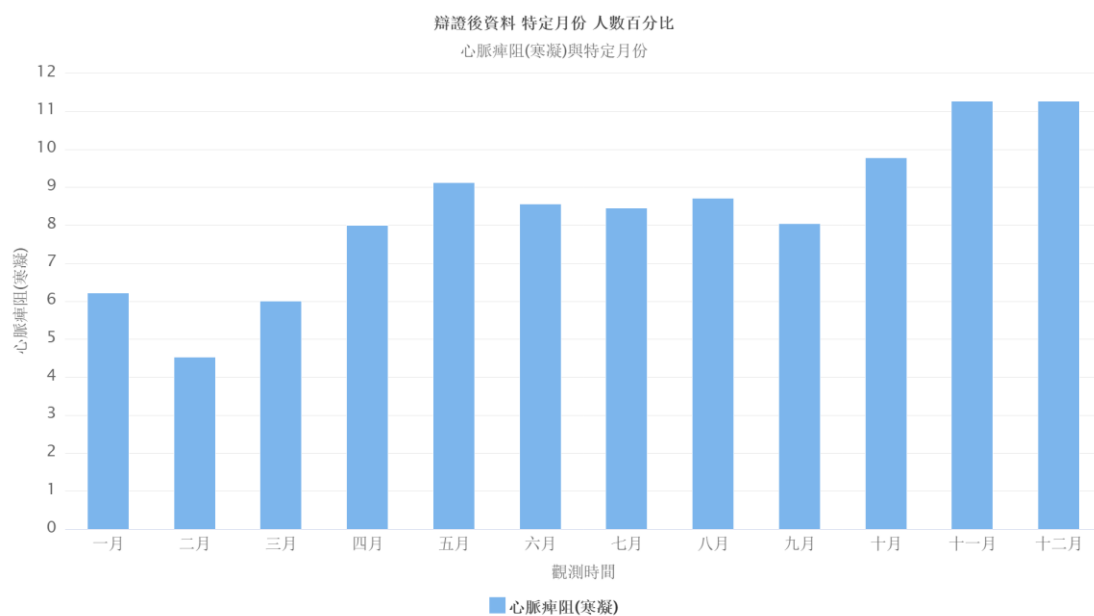


圖 38 心脈痺阻(寒凝)特定月份就診人數百分比長條圖

(圖 39)為心脈痺阻(氣滯)特定月份就診人數百分比長條圖，從圖可看出，心脈痺阻(氣滯)的就診人數百分比在 2 月偏低，而在 8 月偏高。

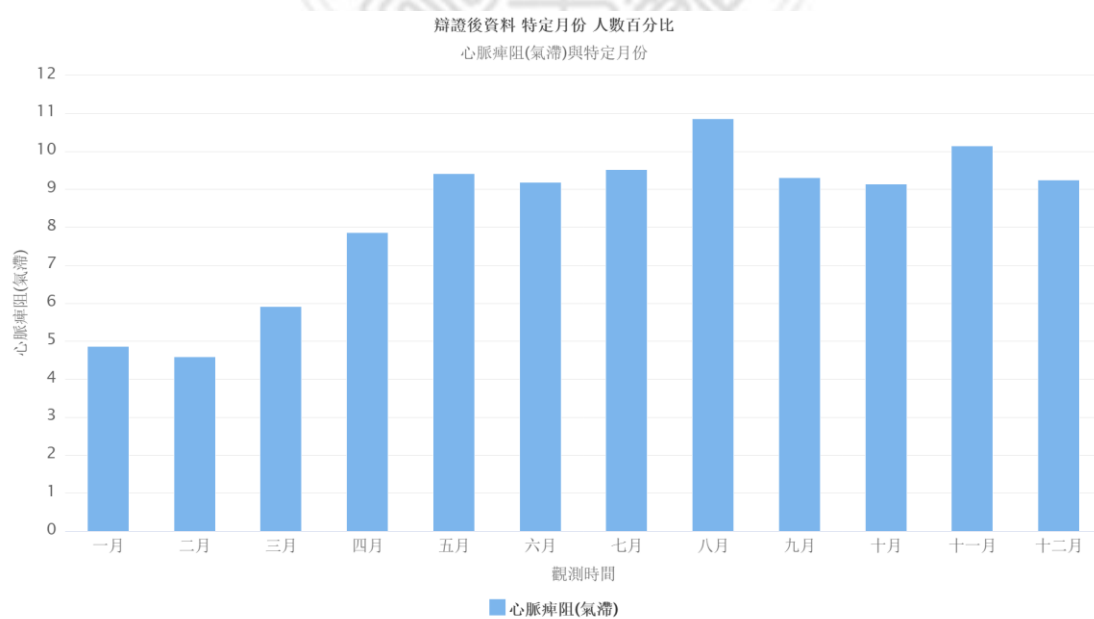


圖 39 心脈痺阻(氣滯)特定月份就診人數百分比長條圖

(圖 40)為肝火上炎特定月份就診人數百分比長條圖，從圖可看出，肝火上炎的就診人數百分比在 2 月偏低，而在 12 月偏高。

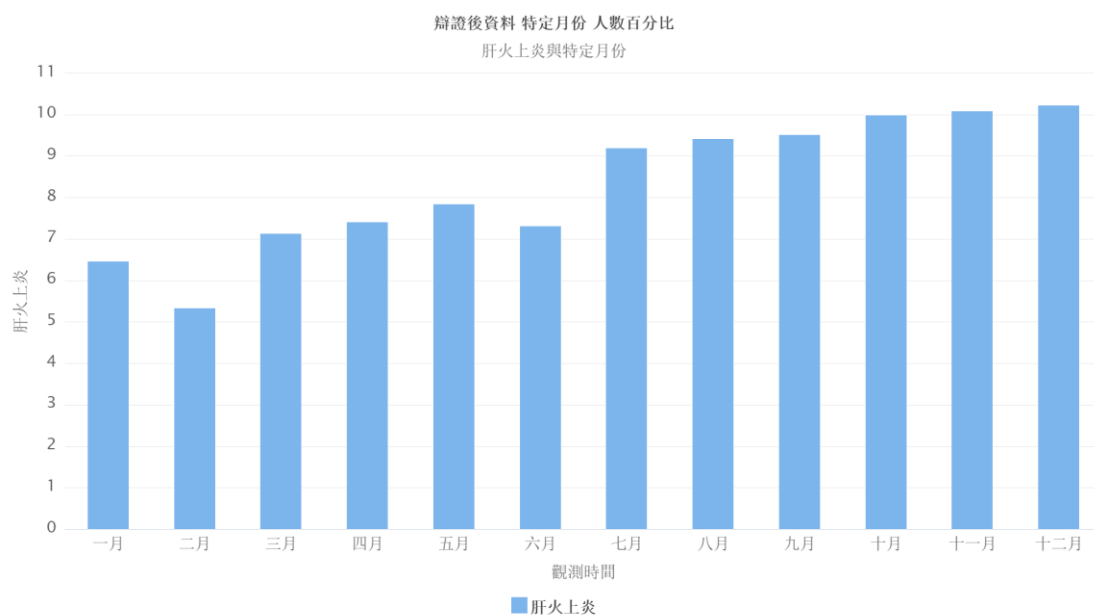


圖 40 肝火上炎特定月份就診人數百分比長條圖

(圖 41)為肝陽上亢特定月份就診人數百分比長條圖，從圖可看出，肝陽上亢的就診人數百分比在 2 月偏低，而 12 月偏高。

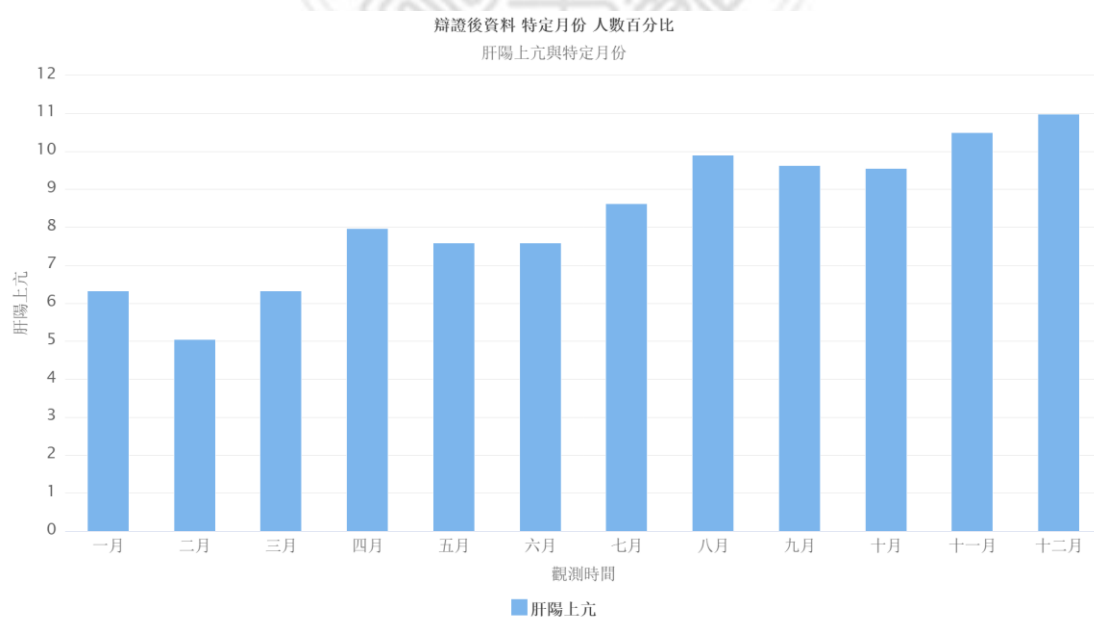


圖 41 肝陽上亢特定月份就診人數百分比長條圖

### 4.1.3. 中醫證候維度與時間維度(每年各季)切割

首先從中醫證候與氣象因子資料方體內切割所需資料，如(圖 42)與(圖 43-44)所示，以資料方體切片之方式依序切出中醫證候之各個階層與時間維度之每月各季階層，其中每年各季之分佈分別為1、2、3月為春季，4、5、6月為夏季，7、8、9月為秋季，10、11、12月為冬季，透過觀察中醫證候與每年各季之資料，找出中醫證候在每年各季不同的關係。

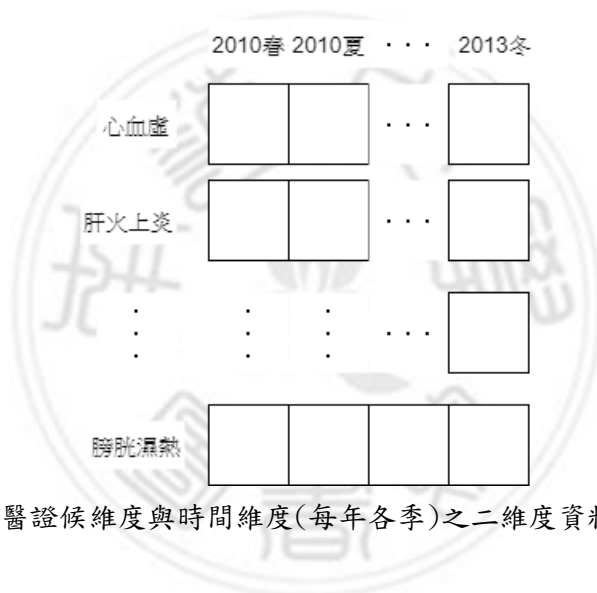


圖 42 中醫證候維度與時間維度(每年各季)之二維度資料方體切割

觀測時間	2010春	2010夏	2010秋	2010冬	2011春	2011夏	2011秋	2011冬	2012春	2012夏	2012秋	2012冬	2013春	2013夏	2013秋	2013冬
心氣虛	78	76	91	93	109	175	177	189	139	173	168	227	248	361	402	405
心血虛	72	74	114	144	116	193	215	228	232	297	303	297	334	408	484	558
心陰虛	72	107	127	165	125	189	181	228	159	240	268	278	254	320	351	394
心陽虛	2	5	10	17	5	37	52	33	21	39	49	54	52	81	46	69
肝血虛	12	15	18	45	43	29	42	41	56	92	62	79	106	108	128	122
脾氣虛	105	152	164	217	169	244	303	405	316	425	342	397	366	495	487	482
脾陽虛	3	14	14	20	8	20	38	24	42	55	26	23	45	53	76	111
脾氣陷	26	28	25	57	54	93	68	79	112	95	112	125	128	162	205	230
胃陰虛	75	127	120	195	152	171	183	180	159	200	222	306	313	311	345	353
肺氣虛	82	85	97	110	109	156	167	186	162	150	155	187	257	316	294	291
肺陰虛	13	26	26	48	38	59	42	30	30	33	53	56	70	120	139	164
腎精不足	20	21	23	28	21	21	16	20	9	29	40	34	28	61	72	78
腎不納氣	63	74	71	64	74	96	127	145	99	141	130	150	164	250	215	197
腎氣不固	113	145	201	238	184	192	189	229	155	207	198	233	258	355	317	280
腎陰虛	40	87	102	146	99	172	163	158	118	188	246	269	223	319	350	302
腎陽虛	70	78	136	168	142	188	177	215	162	234	234	238	250	345	337	309
大腸陽虛	20	31	36	49	60	67	71	109	81	101	108	83	115	116	151	225

圖 43 中醫證候(虛證)與時間維度(每年各季)之資料切片結果

觀測時間	2010春	2010夏	2010秋	2010冬	2011春	2011夏	2011秋	2011冬	2012春	2012夏	2012秋	2012冬	2013春	2013夏	2013秋	2013冬
寒襲衛表	26	44	34	53	54	96	75	117	81	86	118	131	134	144	163	208
暑襲衛表	41	46	48	98	74	93	86	125	119	105	109	166	143	241	159	118
燥襲衛表	27	38	32	83	76	89	66	79	93	87	84	152	136	189	152	124
熱襲衛表	26	41	51	85	84	68	77	98	79	99	121	140	133	210	135	165
心脈痺阻(寒凝)	27	34	42	60	52	135	152	147	96	137	99	129	154	198	202	298
心火亢盛	25	49	29	73	58	73	83	102	110	112	134	149	143	131	150	155
心脈痺阻(痰聚)	65	68	111	115	117	184	205	235	184	234	252	273	263	334	331	436
痰火擾心(鬱)	52	74	80	155	111	92	118	125	169	167	160	206	165	161	205	237
痰火擾心(躁)	27	29	34	26	27	20	14	20	41	46	35	69	56	29	78	84
痰火擾心(外感熱)	65	110	113	144	112	149	166	147	182	214	218	274	298	319	314	280
心脈痺阻(血瘀)	1	11	10	7	17	28	39	39	19	36	30	27	51	98	87	156
心脈痺阻(氣滯)	41	43	54	27	39	114	132	100	73	116	108	112	125	205	242	276
寒邪客肺	10	10	17	29	20	37	20	42	57	46	24	26	40	43	53	24
痰濕阻肺	23	27	48	48	49	55	56	78	101	85	79	87	76	140	153	128
痰熱壅肺	24	20	28	36	35	43	42	55	53	63	43	52	51	98	94	84
大腸液虧	72	98	76	114	93	172	168	128	125	168	195	276	239	226	283	311
大腸實熱	25	38	48	82	62	73	93	102	114	123	149	176	181	126	152	146
寒邪客脾	10	8	7	21	16	17	19	19	34	29	14	27	27	53	86	82
熱邪客脾	31	40	50	66	58	78	73	86	93	145	158	97	118	145	145	146
胃氣上逆	28	47	55	83	73	65	80	86	93	149	103	118	145	199	161	153
食滯胃脘	16	21	26	40	23	44	28	34	44	45	47	44	68	47	77	69
寒滯肝脈	23	30	30	54	43	78	77	86	71	78	67	81	85	100	76	87
肝火上炎	68	116	120	163	144	200	236	233	225	278	365	377	357	353	459	498
肝陽上亢	53	92	141	199	146	196	216	197	209	261	344	325	297	372	419	514
肝氣鬱結	1	4	3	6	13	27	39	30	19	22	18	32	19	48	38	43
膽郁痰擾	51	64	95	131	134	162	128	169	202	260	210	310	252	279	353	343

圖 44 中醫證候(實證)與時間維度(每年各季)之資料切片結果



(圖 45)為心陰虛每年各季就診人數百分比長條圖，從圖可看出，心陰虛的就診人數百分比在每年冬季偏高，而每年春季則偏低。

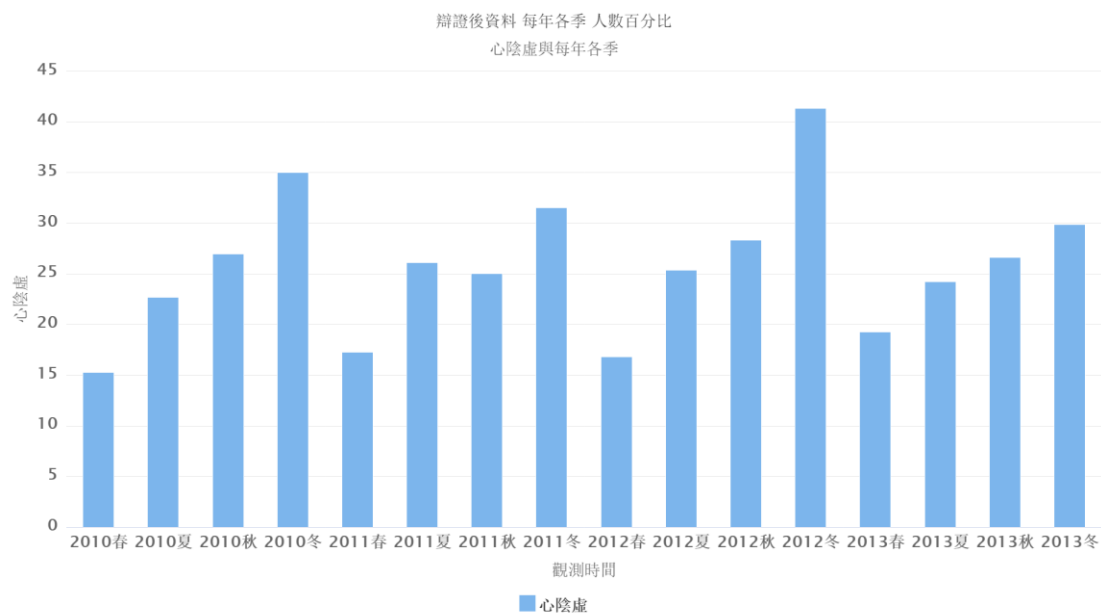


圖 45 心陰虛每年各季就診人數百分比長條圖

(圖 46)為腎不納氣每年各季就診人數百分比長條圖，從圖可看出，腎不納氣的就診人數百分比在 2010、2013 年夏季與 2011、2012 年冬季偏高。



圖 46 腎不納氣每年各季就診人數百分比長條圖

(圖 47)為心脈痺阻(寒凝)每年各季就診人數百分比長條圖，從圖可看出，心脈痺阻(寒凝)的就診人數百分比除 2011 年外，其餘冬季皆偏高，而每年春季皆偏低。

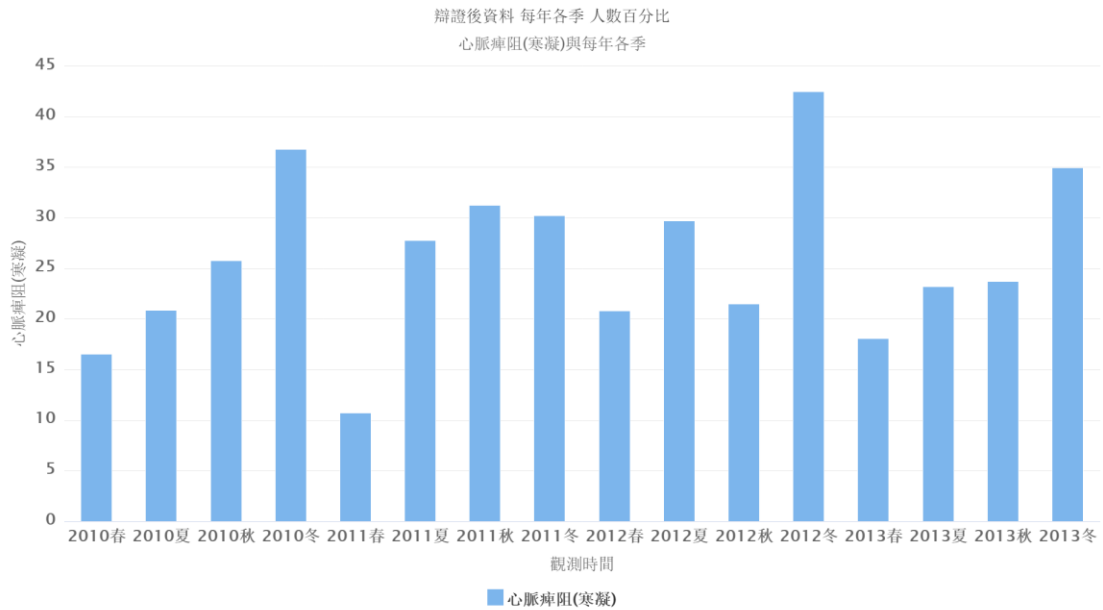


圖 47 心脈痺阻(寒凝)每年各季就診人數百分比長條圖

(圖 48)為心脈痺阻(氣滯)每年各季就診人數百分比長條圖，從圖可看出，心脈痺阻(氣滯)的就診人數百分比在 2010-2011 年秋季與 2012-2013 年冬季皆偏高，而每年春季則皆偏低。



圖 48 心脈痺阻(氣滯)每年各季就診人數百分比長條圖

(圖 49)為肝火上炎每年各季就診人數百分比長條圖，從圖可看出，肝火上炎的就診人數百分比在每年冬季偏高，而每年春季則偏低。

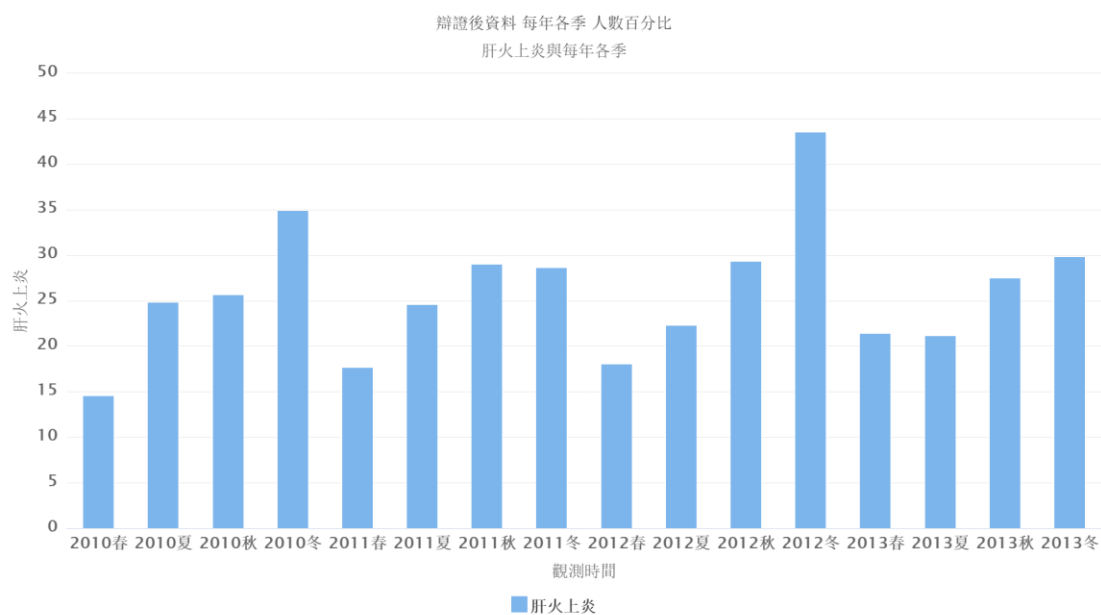


圖 49 肝火上炎每年各季就診人數百分比長條圖

(圖 50)為肝陽上亢每年各季就診人數百分比長條圖，從圖可看出，肝陽上亢的就診人數百分比在除 2011 年外，其餘年冬季偏高，而每年春季則偏低。



圖 50 肝陽上亢每年各季就診人數百分比長條圖

#### 4.1.4. 中醫證候維度與時間維度(特定季節)切割

首先從中醫證候與氣象因子資料方體內切割所需資料，如(圖 51)與(圖 52-53)所示，以資料方體切片之方式依序切出中醫證候之各個階層與時間維度之特定季節階層，其中特定季節為將 2010-2013 年各季總和，透過觀察中醫證候與特定季節之資料，找出中醫證候在特定季節不同的關係。

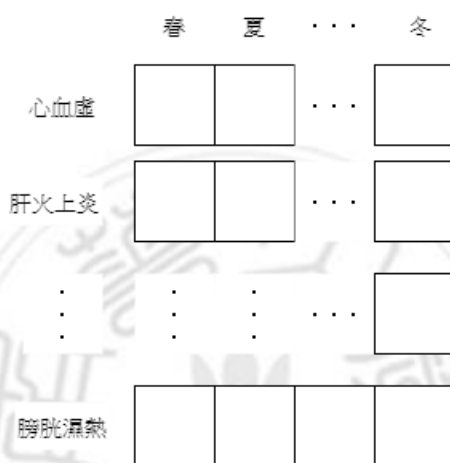


圖 51 中醫證候維度與時間維度(特定季節)切割

觀測時間	春	夏	秋	冬
心氣虛	574	785	838	914
心血虛	754	972	1116	1227
心陰虛	610	856	927	1065
心陽虛	80	162	157	173
肝血虛	217	244	250	287
脾氣虛	956	1316	1296	1501
脾陽虛	98	142	154	178
脾氣陷	320	378	410	491
胃陰虛	699	809	870	1034
肺氣虛	610	707	713	774
肺陰虛	151	238	260	298
腎精不足	78	132	151	160
腎不納氣	400	561	543	556
腎氣不固	710	899	905	980
腎陰虛	480	766	861	875
腎陽虛	624	845	884	930
大腸陽虛	276	315	366	466

圖 52 中醫證候(虛證)與時間維度(特定季節)之資料切片結果

觀測時間	春	夏	秋	冬
寒襲衛表	295	370	390	509
暑襲衛表	377	485	402	507
燥襲衛表	332	403	334	438
熱襲衛表	322	418	384	488
心脈痺阻(寒凝)	329	504	495	634
心火亢盛	336	365	396	479
心脈痺阻(痰聚)	629	820	899	1059
痰火擾心(鬱)	497	494	563	723
痰火擾心(躁)	151	124	161	199
痰火擾心(外感熱)	657	792	811	845
心脈痺阻(血瘀)	88	173	166	229
心脈痺阻(氣滯)	278	478	536	515
寒邪客肺	127	136	114	121
痰濕阻肺	249	307	336	341
痰熱壅肺	163	224	207	227
大腸液虧	529	664	722	829
大腸實熱	382	360	442	506
寒邪客脾	87	107	126	149
熱邪客脾	300	408	426	395
胃氣上逆	339	460	399	440
食滯胃脘	151	157	178	187
寒滯肝脈	222	286	250	308
肝火上炎	794	947	1180	1271
肝陽上亢	705	921	1120	1235
肝氣鬱結	52	101	98	111
膽郁痰擾	639	765	786	953

圖 53 中醫證候(實證)與時間維度(特定季節)之資料切片結果

(圖 54)為腎不納氣特定季節就診人數百分比長條圖，從圖可看出，腎不納氣的就診人數百分比在夏季偏高，而春季則偏低。

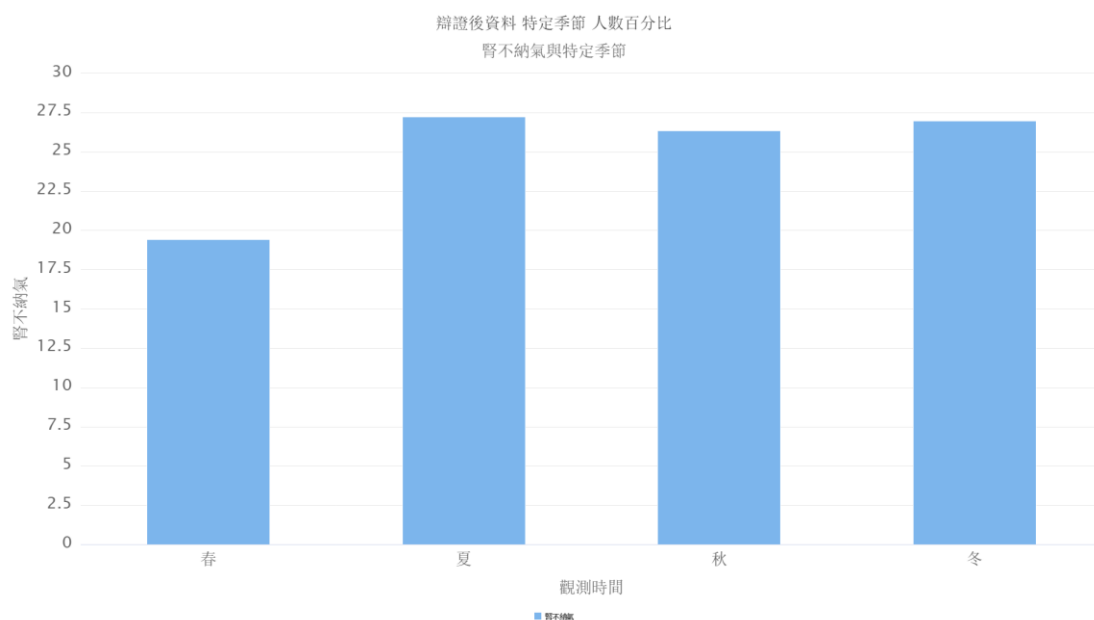


圖 54 腎不納氣特定季節就診人數百分比長條圖

(圖 55)為暑襲衛表特定季節就診人數百分比長條圖，從圖可看出，暑襲衛表的就診人數百分比在冬季偏高，而春季則偏低。

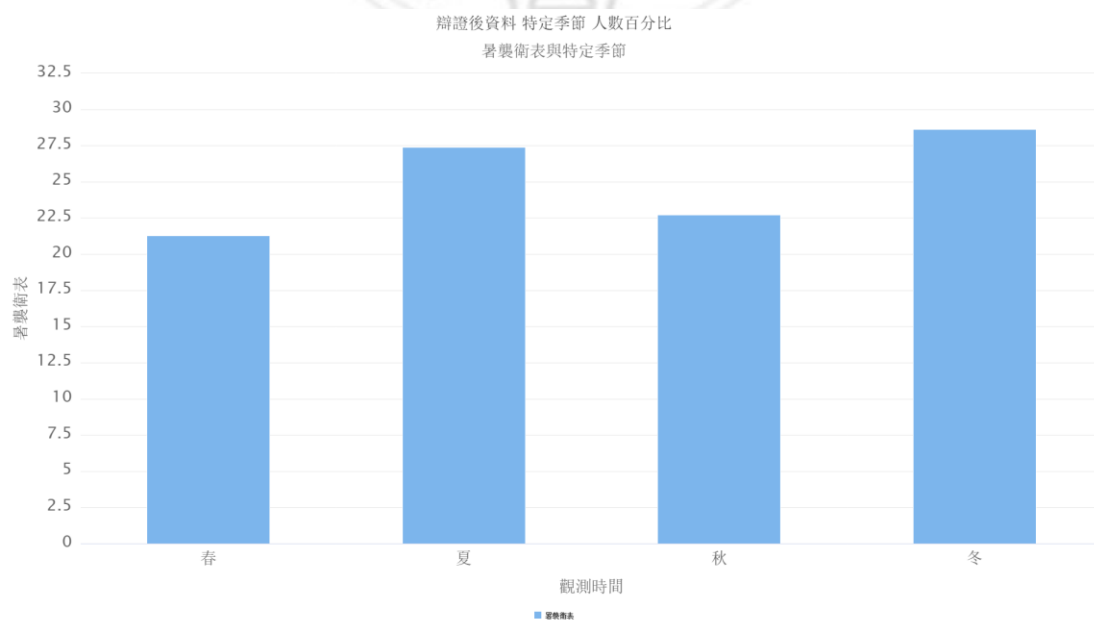


圖 55 暑襲衛表特定季節就診人數百分比長條圖

(圖 56)為心脈痺阻(寒凝)特定季節就診人數百分比長條圖，從圖可看出，心脈痺阻(寒凝)的就診人數百分比在冬季偏高，而春季則偏低。

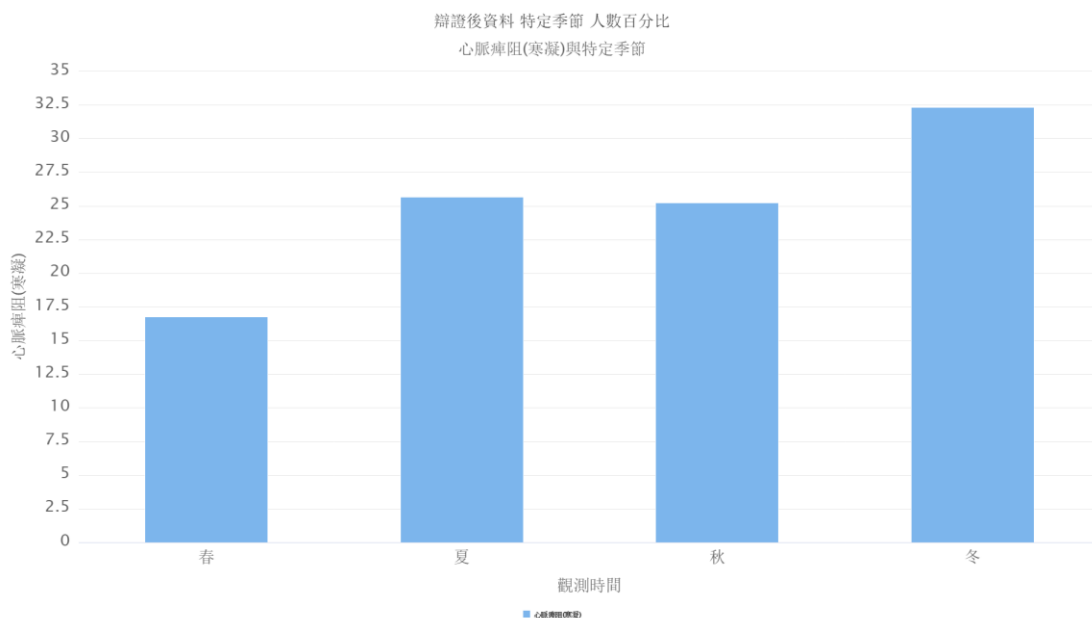


圖 56 心脈痺阻(寒凝)特定季節就診人數百分比長條圖

(圖 57)為心脈痺阻(氣滯)特定季節就診人數百分比長條圖，從圖可看出，心脈痺阻(氣滯)的就診人數百分比在秋季偏高，而春季則偏低。

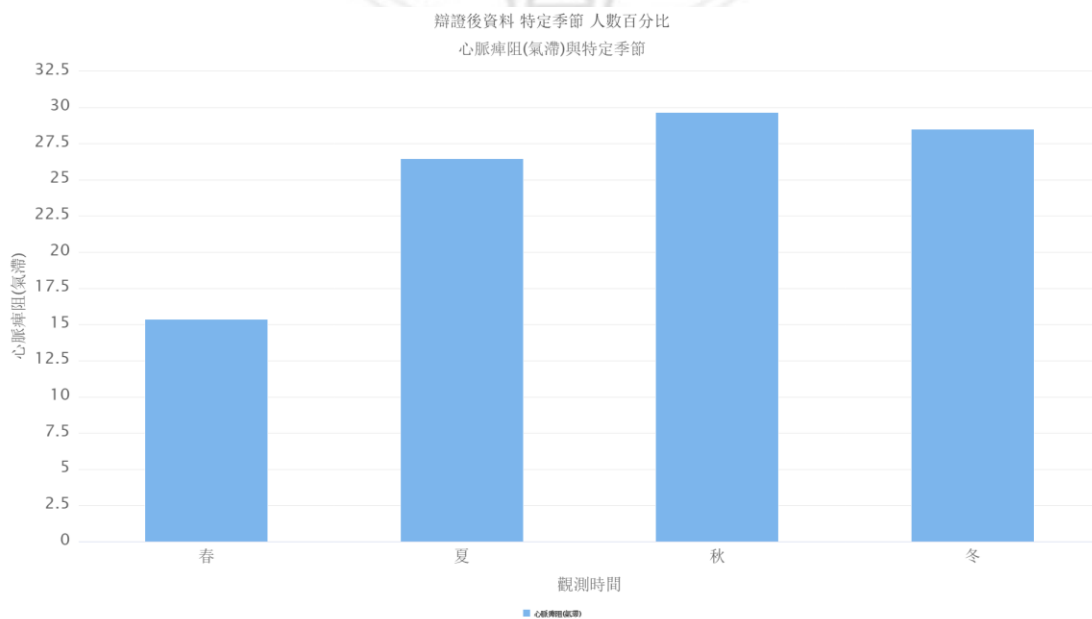


圖 57 心脈痺阻(氣滯)特定季節就診人數百分比長條圖

(圖 58)為肝火上炎特定季節就診人數百分比長條圖，從圖可看出，肝火上炎的就診人數百分比在冬季偏高，而春季則偏低。

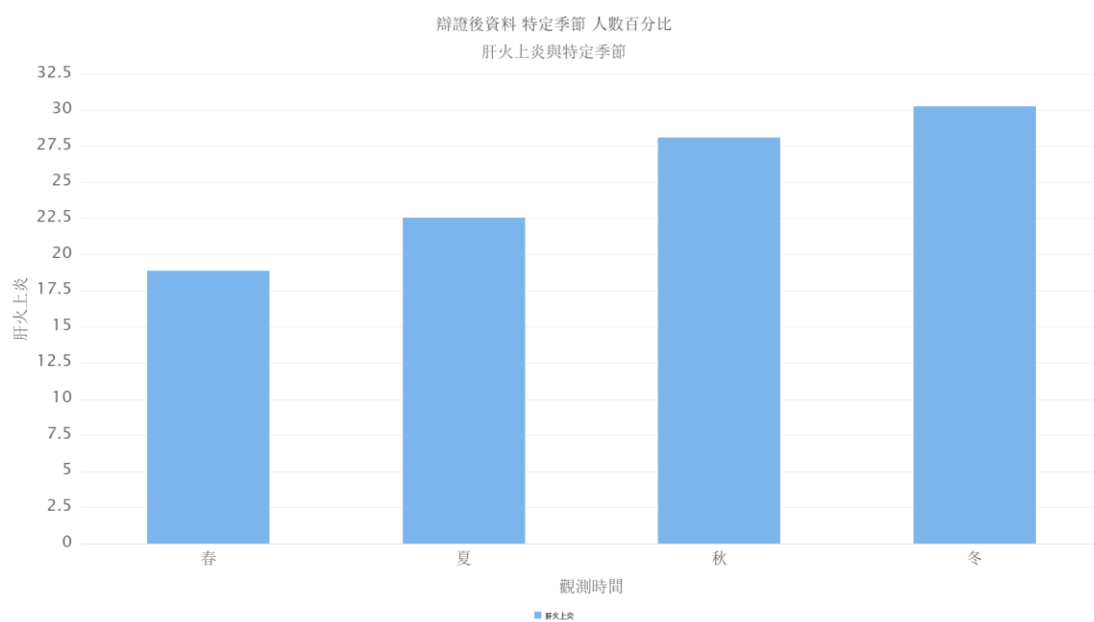


圖 58 肝火上炎特定季節就診人數百分比長條圖

(圖 59)為肝陽上亢特定季節就診人數百分比長條圖，從圖可看出，肝陽上亢的就診人數百分比在冬季偏高，而春季則偏低。

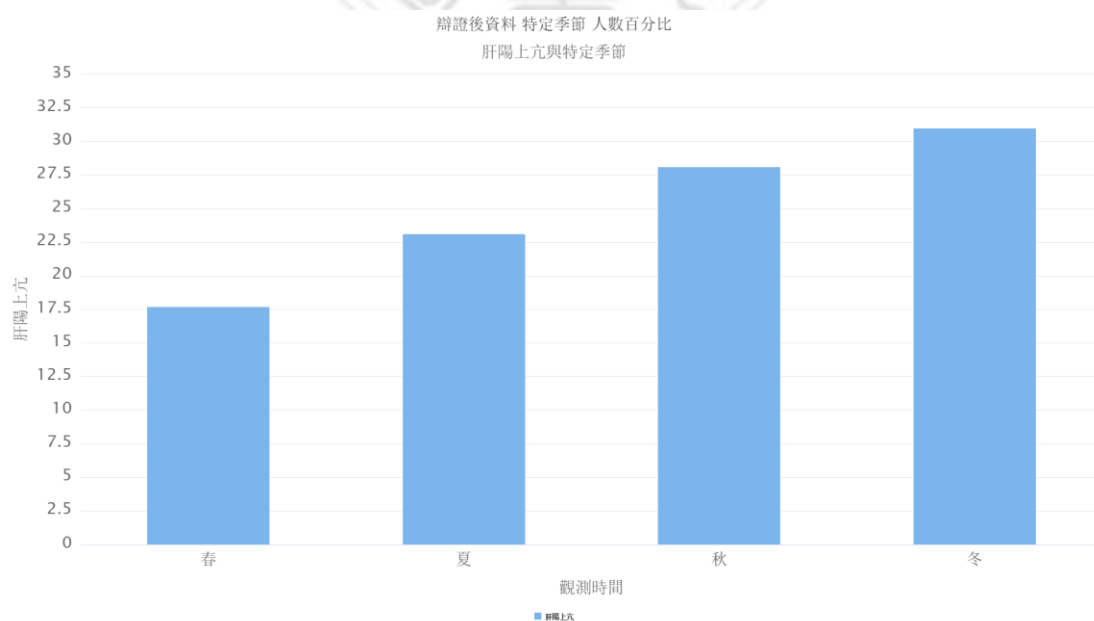


圖 59 肝陽上亢特定季節就診人數百分比長條圖



#### 4.1.5. 中醫證候維度與時間維度(每年六氣)切割

首先從中醫證候與氣象因子資料方體內切割所需資料，如(圖 60)與(圖 61-66)所示，但由於 2013 年之寒證並無資料，故採用 2010-2012 年資料進行觀察，先以資料方體切片之方式依序切出中醫證候之各個階層與時間維度之每年六氣階層，六氣時段依據二十四節氣進行分類，每一氣包含了四個節氣(胡建勳、劉堯欽，2008)，如(表 五)，透過觀察中醫證候與每年六氣之資料，找出中醫證候在每年六氣不同的關係。

表 五 六氣時段表

六氣	包含之二十四節氣	日期區間
風	大寒、立春、雨水、驚蟄	1/20 - 3/19
暑	春分、清明、穀雨、立夏	3/20 - 5/20
火	小滿、芒種、夏至、小暑	5/21 - 7/22
濕	大暑、立秋、處暑、白露	7/23 - 9/22
燥	秋分、寒露、霜降、立冬	9/23 - 11/21
寒	小雪、大雪、冬至、小寒	11/22 - 1/19

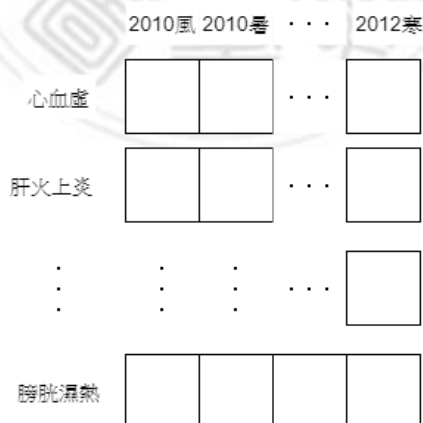


圖 60 中醫證候維度與時間維度(每年六氣)之二維度資料方體切割

觀測時間	2010風	2010暑	2010火	2010濕	2010燥	2010寒
心氣虛	54	60	47	68	46	70
心血虛	41	52	58	84	86	88
心陰虛	45	76	66	97	109	96
心陽虛	2	3	3	8	9	9
肝血虛	11	10	12	9	22	38
脾氣虛	72	105	94	110	130	146
脾陽虛	2	5	12	9	14	8
脾氣陷	14	22	21	13	35	34
胃陰虛	40	98	74	85	122	111
肺氣虛	55	64	49	71	58	76
肺陰虛	5	16	17	20	34	27
腎精不足	12	20	9	16	20	14
腎不納氣	41	54	47	50	39	39
腎氣不固	81	95	104	139	141	159
腎陰虛	27	51	57	73	97	87
腎陽虛	49	52	70	92	105	100
大腸陽虛	13	18	22	27	24	36

圖 61 中醫證候(虛證)與時間維度(2010年六氣)之資料切片結果

觀測時間	2010風	2010暑	2010火	2010濕	2010燥	2010寒
寒襲衛表	17	30	28	22	26	43
暑襲衛表	27	32	28	34	53	64
燥襲衛表	15	29	19	22	45	54
熱襲衛表	18	23	32	34	39	68
心脈痺阻(寒凝)	17	25	18	34	42	34
心火亢盛	12	38	23	22	47	45
心脈痺阻(痰聚)	41	54	44	90	62	87
痰火擾心(鬱)	27	53	46	62	84	106
痰火擾心(躁)	15	19	27	19	22	16
痰火擾心(外感熱)	38	79	72	75	108	72
心脈痺阻(血瘀)	0	7	6	8	4	4
心脈痺阻(氣滯)	25	32	28	43	17	16
寒邪客肺	5	8	5	14	21	9
痰濕阻肺	15	18	17	41	24	35
痰熱壅肺	18	15	10	24	17	26
大腸液虧	43	71	68	46	77	60
大腸實熱	14	28	24	33	57	44
寒邪客脾	7	6	7	4	8	13
熱邪客脾	23	30	19	36	33	55
胃氣上逆	17	36	23	41	43	60
食滯胃脘	11	14	19	14	27	19
寒滯肝脈	16	20	17	23	35	32
肝火上炎	36	87	79	75	103	102
肝陽上亢	31	61	65	102	115	133
肝氣鬱結	1	3	1	2	5	2
膽郁痰擾	27	51	43	70	74	94

圖 62 中醫證候(實證)與時間維度(每年六氣)之資料切片結果

觀測時間	2011風	2011暑	2011火	2011濕	2011燥	2011寒
心氣虛	66	122	117	114	122	126
心血虛	68	135	139	134	159	152
心陰虛	75	124	132	115	149	140
心陽虛	4	21	29	35	26	23
肝血虛	24	24	19	28	38	22
脾氣虛	106	151	198	187	245	276
脾陽虛	7	10	26	19	18	19
脾氣陷	36	55	69	41	48	54
胃陰虛	105	122	110	127	120	109
肺氣虛	70	105	114	105	113	133
肺陰虛	18	47	31	28	23	23
腎精不足	14	17	10	12	15	8
腎不納氣	46	70	80	77	101	85
腎氣不固	112	129	143	115	145	157
腎陰虛	60	111	115	113	100	98
腎陽虛	98	129	124	116	121	145
大腸陽虛	47	48	50	41	57	77

圖 63 中醫證候(虛證)與時間維度(2011年六氣)之資料切片結果

觀測時間	2011風	2011暑	2011火	2011濕	2011燥	2011寒
寒襲衛表	36	61	60	48	63	76
暑襲衛表	51	60	61	55	70	94
燥襲衛表	56	66	48	42	45	62
熱襲衛表	52	60	35	55	70	55
心脈痺阻(寒凝)	34	81	97	103	89	101
心火亢盛	31	49	60	47	68	64
心脈痺阻(痰聚)	71	118	144	124	135	162
痰火擾心(鬱)	66	77	59	79	83	94
痰火擾心(躁)	17	11	16	7	13	17
痰火擾心(外感熱)	75	87	127	100	98	101
心脈痺阻(血瘀)	16	17	21	29	18	25
心脈痺阻(氣滯)	27	68	83	93	64	64
寒邪客肺	15	30	15	14	17	39
痰濕阻肺	27	53	28	40	43	52
痰熱壅肺	23	38	24	28	28	38
大腸液虧	54	129	107	111	87	86
大腸實熱	38	50	61	57	68	64
寒邪客脾	15	9	15	12	10	18
熱邪客脾	40	52	60	42	44	62
胃氣上逆	53	53	38	56	49	61
食滯胃脘	15	30	27	17	20	22
寒滯肝脈	29	56	47	50	47	67
肝火上炎	88	143	149	144	152	163
肝陽上亢	88	141	131	143	128	146
肝氣鬱結	11	19	20	28	20	16
膽郁痰擾	82	121	98	84	109	114

圖 64 中醫證候(實證)與時間維度(2011年六氣)之資料切片結果

觀測時間	2012風	2012暑	2012火	2012濕	2012燥	2012寒
心氣虛	86	98	145	104	130	140
心血虛	149	157	240	207	183	206
心陰虛	96	127	205	175	186	181
心陽虛	6	23	34	34	32	30
肝血虛	39	48	65	39	47	64
脾氣虛	196	251	312	212	275	229
脾陽虛	26	29	45	13	6	24
脾氣陷	77	59	87	71	75	85
胃陰虛	98	137	147	149	183	218
肺氣虛	98	104	118	95	126	104
肺陰虛	14	24	29	35	31	41
腎精不足	5	17	23	27	26	18
腎不納氣	58	81	115	76	105	78
腎氣不固	86	133	150	123	159	128
腎陰虛	71	107	164	155	188	163
腎陽虛	107	130	194	145	150	157
大腸陽虛	52	62	73	74	54	63

圖 65 中醫證候(虛證)與時間維度(2012年六氣)之資料切片結果

觀測時間	2012風	2012暑	2012火	2012濕	2012燥	2012寒
寒襲衛表	54	67	58	83	86	74
暑襲衛表	70	90	61	72	113	88
燥襲衛表	55	68	57	55	116	77
熱襲衛表	48	70	73	79	93	78
心脈痺阻(寒凝)	50	90	99	53	77	98
心火亢盛	74	65	96	91	100	97
心脈痺阻(痰聚)	113	143	200	151	165	189
痰火擾心(鬱)	105	106	127	107	128	128
痰火擾心(躁)	27	28	32	25	45	46
痰火擾心(外感熱)	122	139	147	152	187	188
心脈痺阻(血瘀)	12	26	22	18	13	33
心脈痺阻(氣滯)	44	63	96	69	65	84
寒邪客肺	37	29	35	11	22	16
痰濕阻肺	64	68	62	49	58	51
痰熱壅肺	34	42	44	24	38	29
大腸液虧	81	99	139	123	171	183
大腸實熱	76	72	102	104	119	118
寒邪客脾	19	18	26	6	14	20
熱邪客脾	59	92	116	98	73	61
胃氣上逆	55	91	113	57	73	80
食滯胃脘	31	26	36	33	37	26
寒滯肝脈	38	62	50	38	47	59
肝火上炎	138	178	219	250	243	251
肝陽上亢	129	164	202	238	188	228
肝氣鬱結	11	17	15	11	17	19
膽郁痰擾	131	140	208	131	194	187

圖 66 中醫證候(實證)與時間維度(2012年六氣)之資料切片結果

(圖 67)為腎氣不固每年六氣就診人數長條圖，從圖可看出，腎氣不固的就診人數在每年風證時偏少，而在 2010-2011 年寒證與 2012 年燥證時偏多。

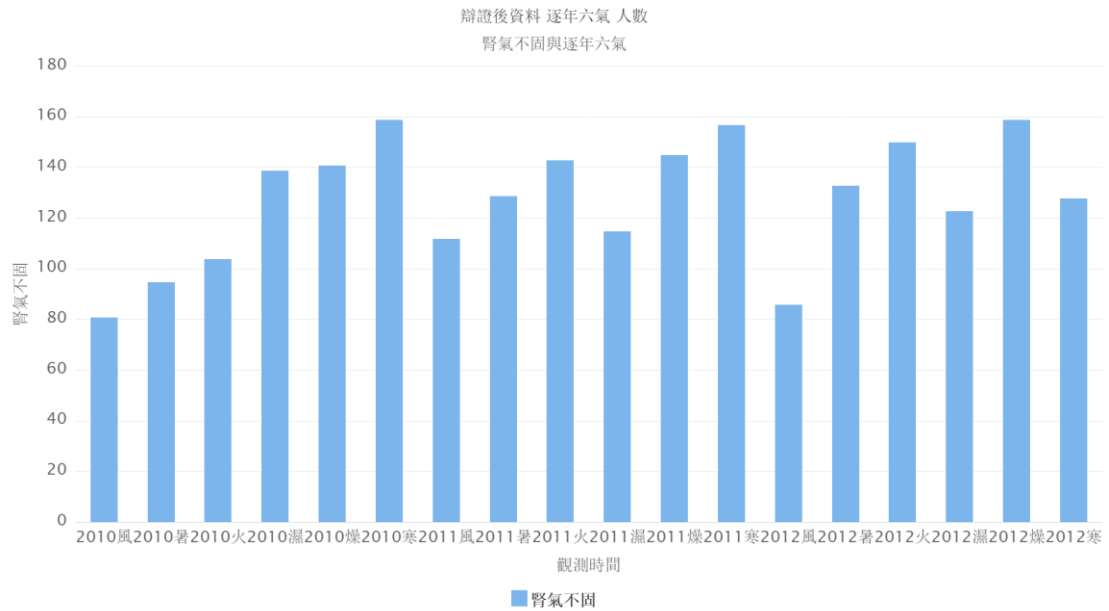


圖 67 腎氣不固每年六氣就診人數長條圖

(圖 68)為寒襲衛表每年六氣就診人數長條圖，從圖可看出，寒襲衛表的就診人數在每年風證時偏少，而在 2010-2011 年寒證與 2012 年燥證時偏多。

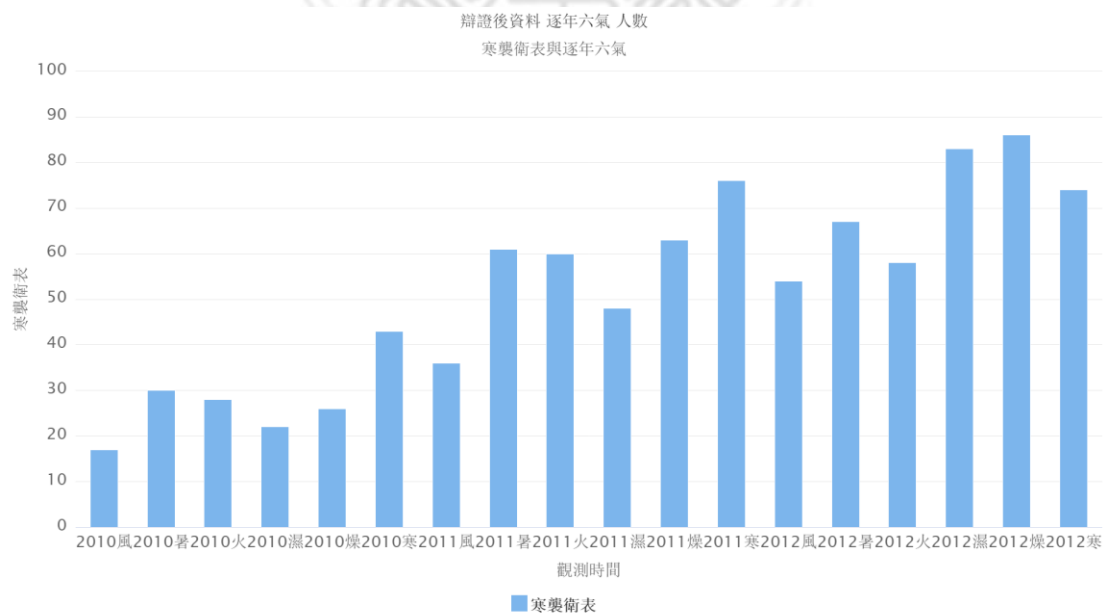


圖 68 寒襲衛表每年六氣就診人數長條圖

(圖 69)為心脈痺阻(寒凝)每年六氣就診人數長條圖，從圖可看出，心脈痺阻(寒凝)的就診人數在每年風證時偏少，而在 2010 年燥、2011 年濕證與 2012 年火證時偏多。

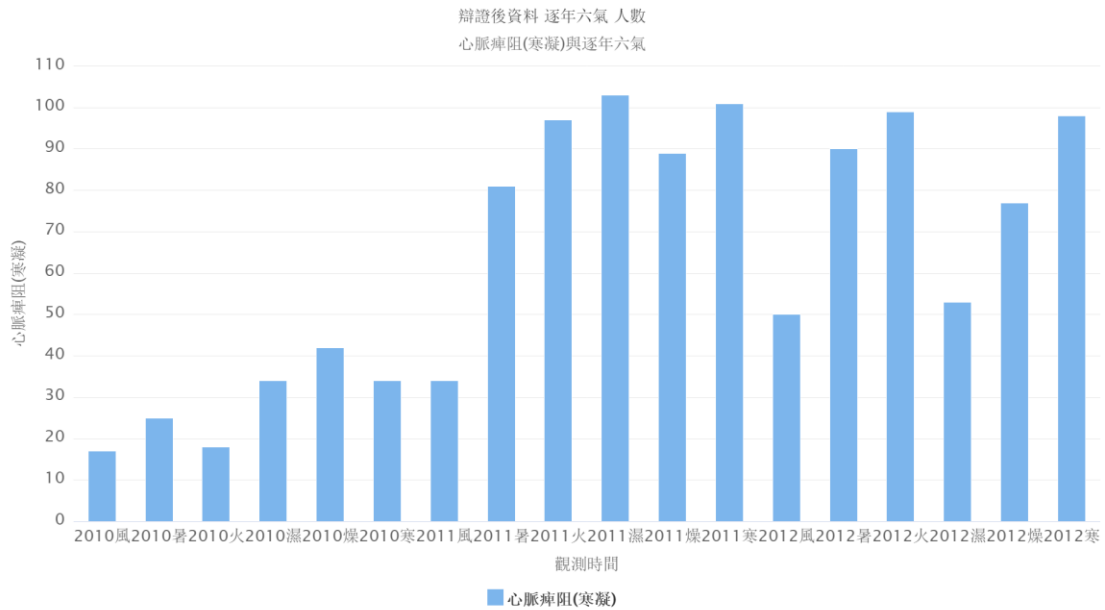


圖 69 心脈痺阻(寒凝)每年六氣就診人數長條圖

(圖 70)為心脈痺阻(氣滯)每年六氣就診人數長條圖，從圖可看出，心脈痺阻(氣滯)的就診人數在 2010 年寒證與 2011-2012 年風證時偏少，2010-2011 年濕證與 2012 年火證時偏多。

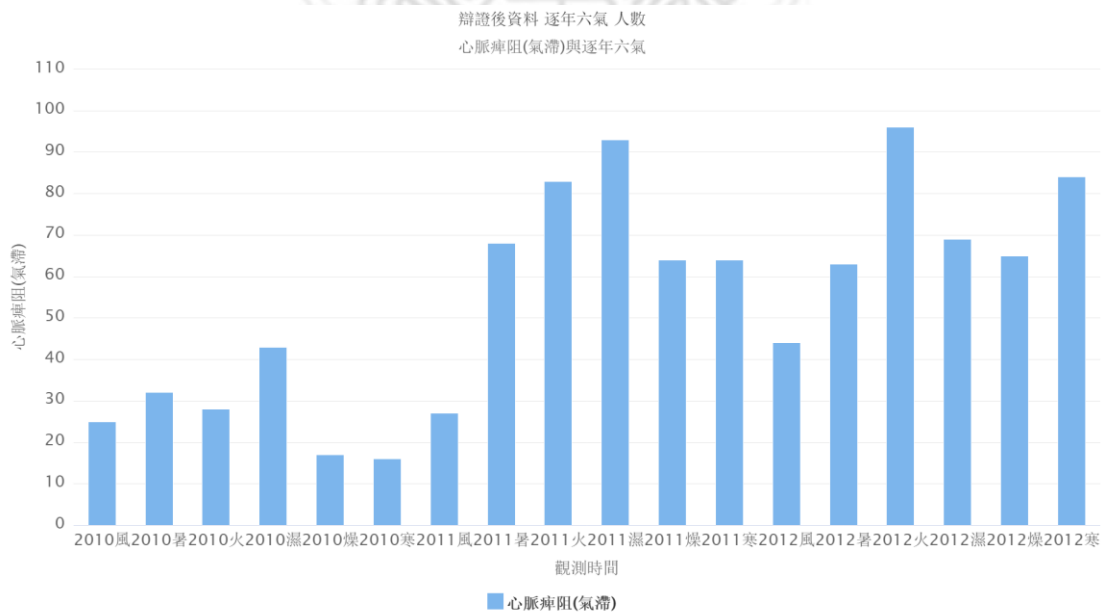


圖 70 心脈痺阻(氣滯)每年六氣就診人數長條圖

(圖 71)為肝火上炎每年六氣就診人數長條圖，從圖可看出，肝火上炎的就診人數在每年風證時偏少，而在每年寒證時偏多。

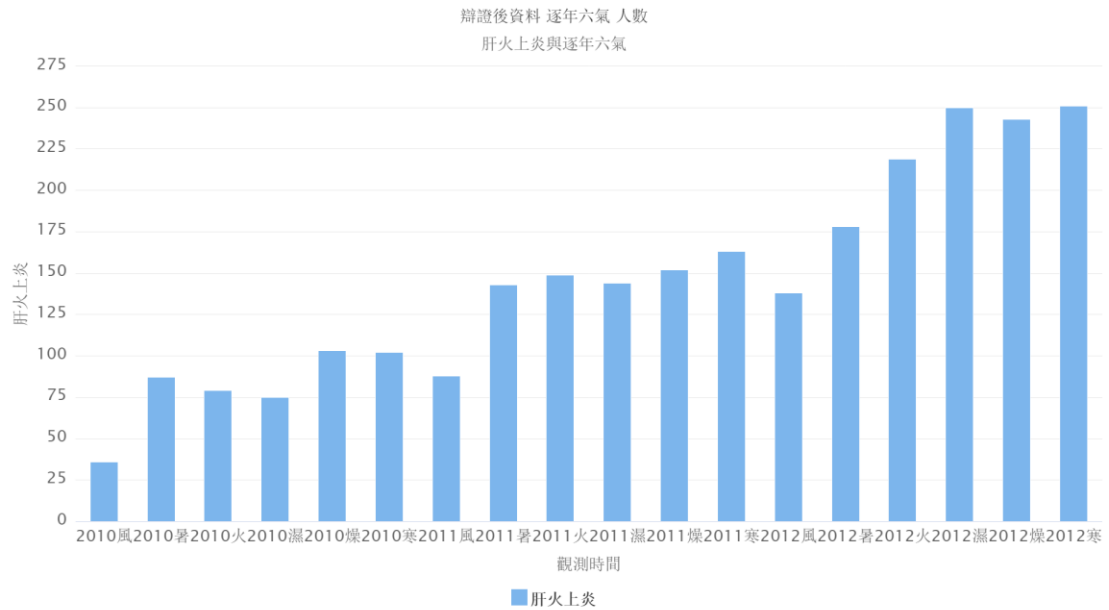


圖 71 肝火上炎每年六氣就診人數長條圖

(圖 72)為肝陽上亢每年六氣就診人數長條圖，從圖可看出，肝陽上亢的就診人數在每年風證時偏少，2010-2011 年寒證與 2012 年濕證時偏多。



圖 72 肝陽上亢每年六氣就診人數長條圖

#### 4.1.6. 中醫證候維度與時間維度(特定六氣)切割

首先從中醫證候與氣象因子資料方體內切割所需資料，如(圖 73)與(圖 74-75)所示，但由於 2013 年之寒證並無資料，故採用 2010-2012 年資料進行觀察，先以資料方體切片之方式依序切出中醫證候之各個階層與時間維度之特定六氣階層透過觀察中醫證候與特定六氣之資料，找出中醫證候與特定六氣不同的關係。

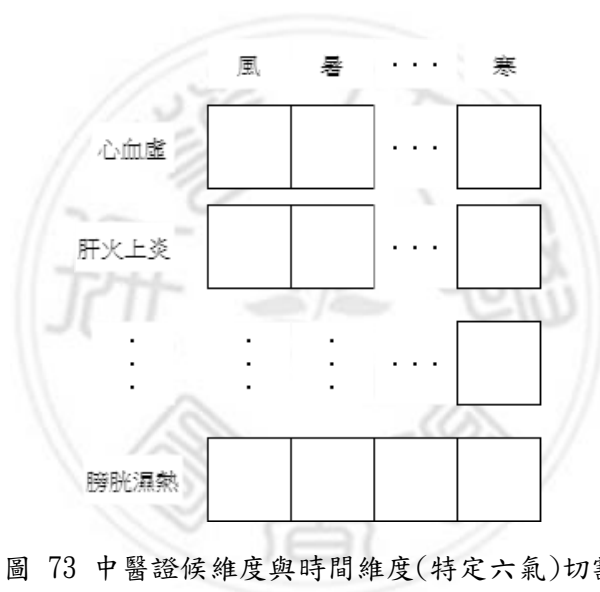


圖 73 中醫證候維度與時間維度(特定六氣)切割



觀測時間	風	暑	火	濕	燥	寒
心氣虛	206	280	309	286	298	336
心血虛	258	344	437	425	428	446
心陰虛	216	327	403	387	444	417
心陽虛	12	47	66	77	67	62
肝血虛	74	82	96	76	107	124
脾氣虛	374	507	604	509	650	651
脾陽虛	35	44	83	41	38	51
脾氣陷	127	136	177	125	158	173
胃陰虛	243	357	331	361	425	438
肺氣虛	223	273	281	271	297	313
肺陰虛	37	87	77	83	88	91
腎精不足	31	54	42	55	61	40
腎不納氣	145	205	242	203	245	202
腎氣不固	279	357	397	377	445	444
腎陰虛	158	269	336	341	385	348
腎陽虛	254	311	388	353	376	402
大腸陽虛	112	128	145	142	135	176

圖 74 中醫證候(虛證)與時間維度(特定六氣)之資料切片結果

觀測時間	風	暑	火	濕	燥	寒
寒襲衛表	107	158	146	153	175	193
暑襲衛表	148	182	150	161	236	246
燥襲衛表	126	163	124	119	206	193
熱襲衛表	118	153	140	168	202	201
心脈痺阻(寒凝)	101	196	214	190	208	233
心火亢盛	117	152	179	160	215	206
心脈痺阻(痰聚)	225	315	388	365	362	438
痰火擾心(鬱)	198	236	232	248	295	328
痰火擾心(躁)	59	58	75	51	80	79
痰火擾心(外感熱)	235	305	346	327	393	361
心脈痺阻(血瘀)	28	50	49	55	35	62
心脈痺阻(氣滯)	96	163	207	205	146	164
寒邪客肺	57	67	55	39	60	64
痰濕阻肺	106	139	107	130	125	138
痰熱壅肺	75	95	78	76	83	93
大腸液虧	178	299	314	280	335	329
大腸實熱	128	150	187	194	244	226
寒邪客脾	41	33	48	22	32	51
熱邪客脾	122	174	195	176	150	178
胃氣上逆	125	180	174	154	165	201
食滯胃脘	57	70	82	64	84	67
寒滯肝脈	83	138	114	111	129	158
肝火上炎	262	408	447	469	498	516
肝陽上亢	248	366	398	483	431	507
肝氣鬱結	23	39	36	41	42	37
膽郁痰擾	240	312	349	285	377	395

圖 75 中醫證候(實證)與時間維度(特定六氣)之資料切片結果

(圖 76)為心陰虛特定六氣就診人數長條圖，從圖可看出，心陰虛的就診人數在風證時偏少，而在燥證時偏多。

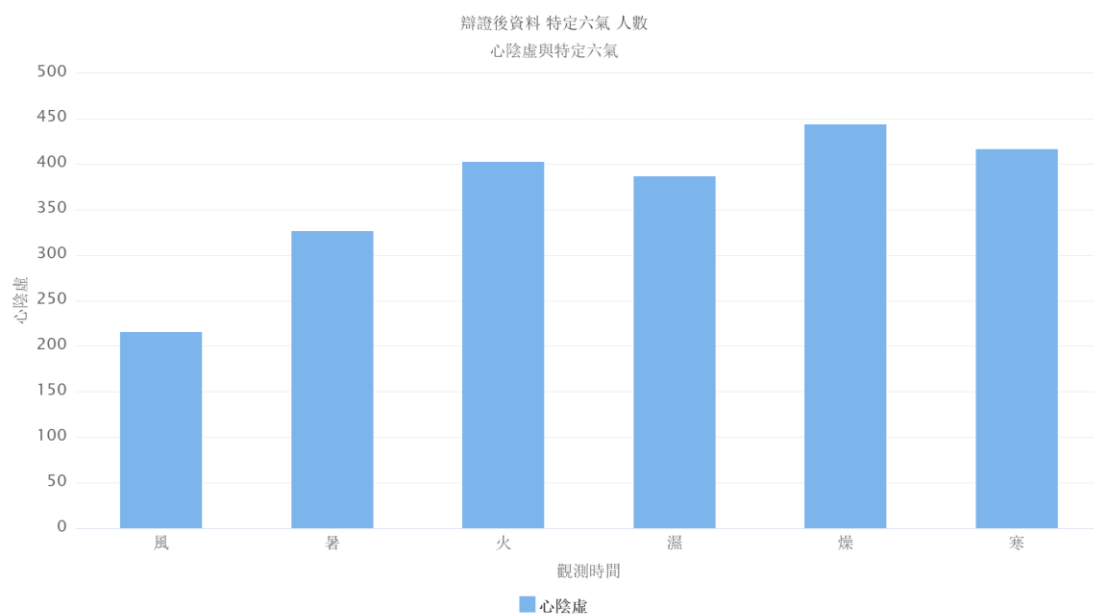


圖 76 心陰虛特定六氣就診人數長條圖

(圖 77)為寒襲衛表特定六氣就診人數長條圖，從圖可看出，寒襲衛表的就診人數在風證時偏少，而在寒證時偏多。

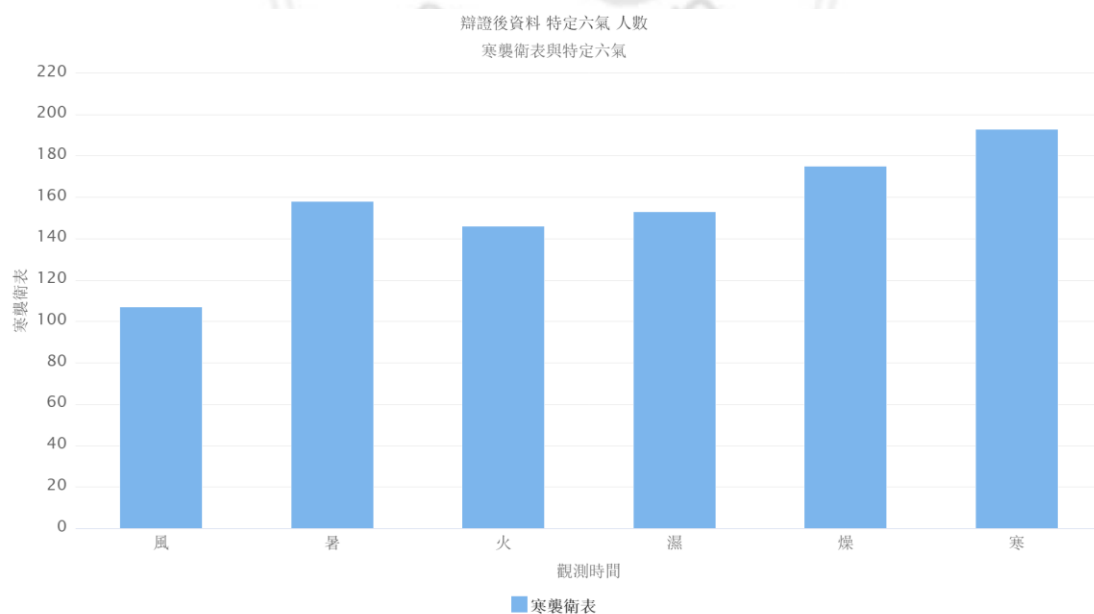


圖 77 寒襲衛表特定六氣就診人數長條圖

(圖 78)為心脈痺阻(寒凝)特定六氣就診人數長條圖，從圖可看出，心脈痺阻(寒凝)的就診人數在風證時偏少，而在寒證時偏多。

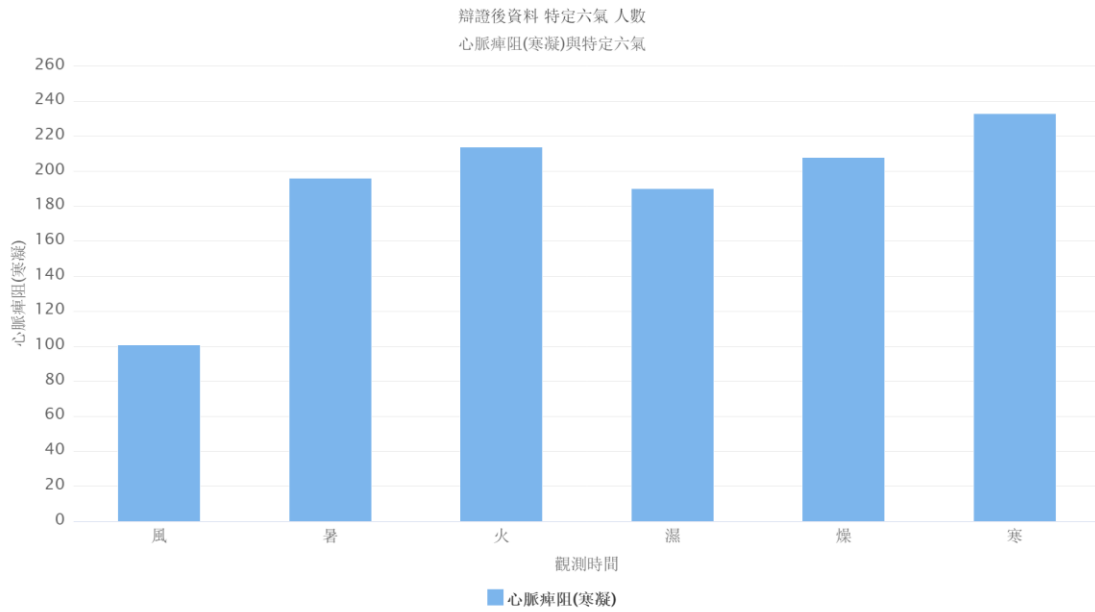


圖 78 心脈痺阻(寒凝)特定六氣就診人數長條圖

(圖 79)為心脈痺阻(氣滯)特定六氣就診人數長條圖，從圖可看出，心脈痺阻(氣滯)的就診人數在風證時偏少，而在火證時偏多。

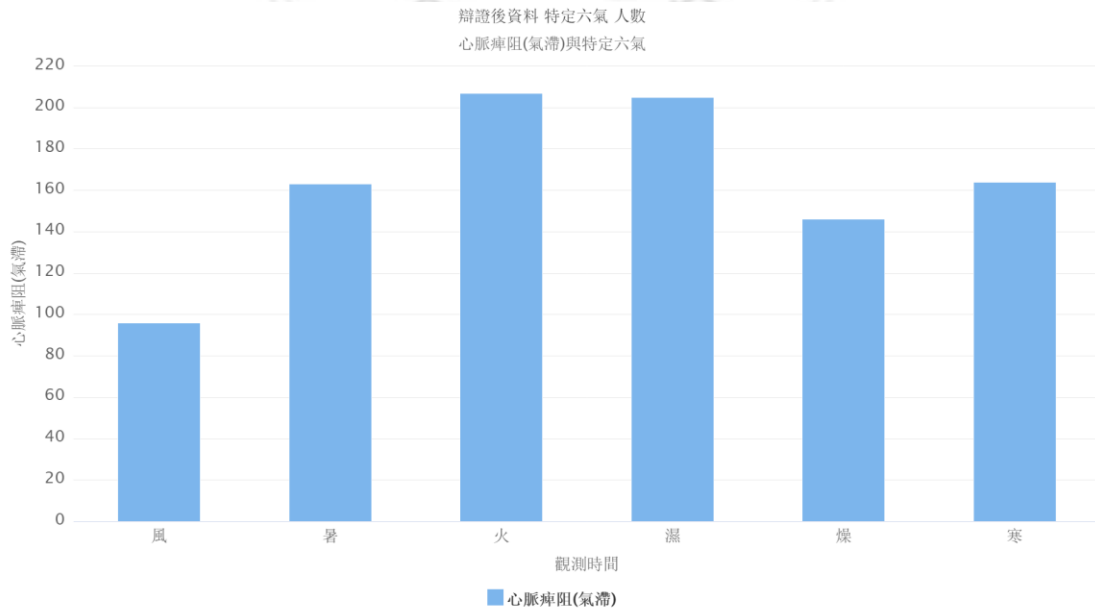


圖 79 心脈痺阻(氣滯)特定六氣就診人數長條圖

(圖 80)為肝火上炎特定六氣就診人數長條圖，從圖可看出，肝火上炎的就診人數在風證時偏少，而在寒證時偏多。

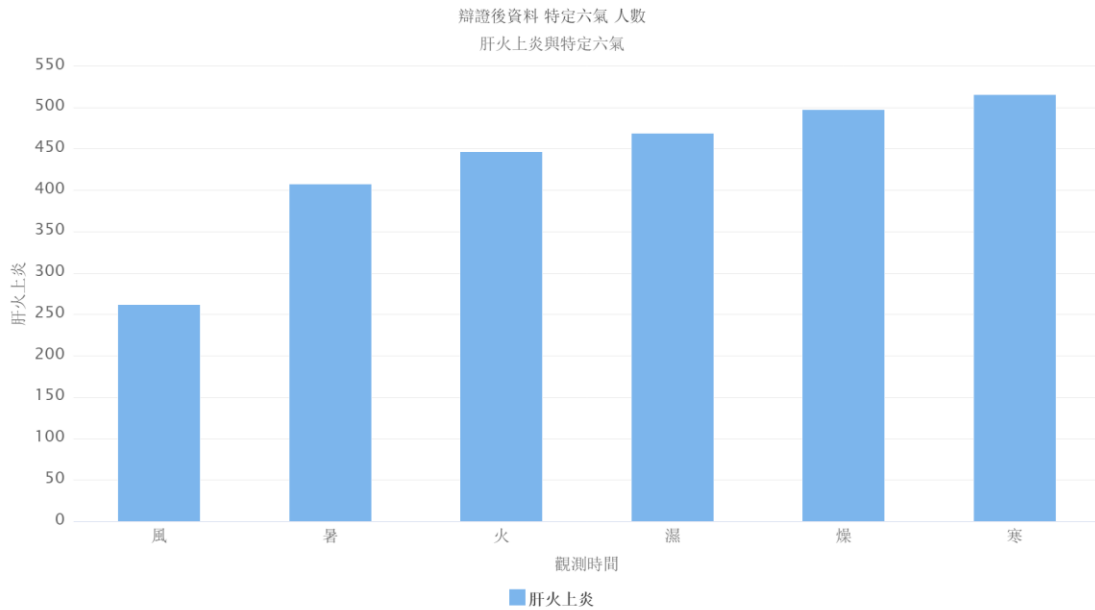


圖 80 肝火上炎特定六氣就診人數長條圖

(圖 81)為肝陽上亢特定六氣就診人數長條圖，從圖可看出，肝陽上亢的就診人數在風證時偏少，而在寒證時偏多。

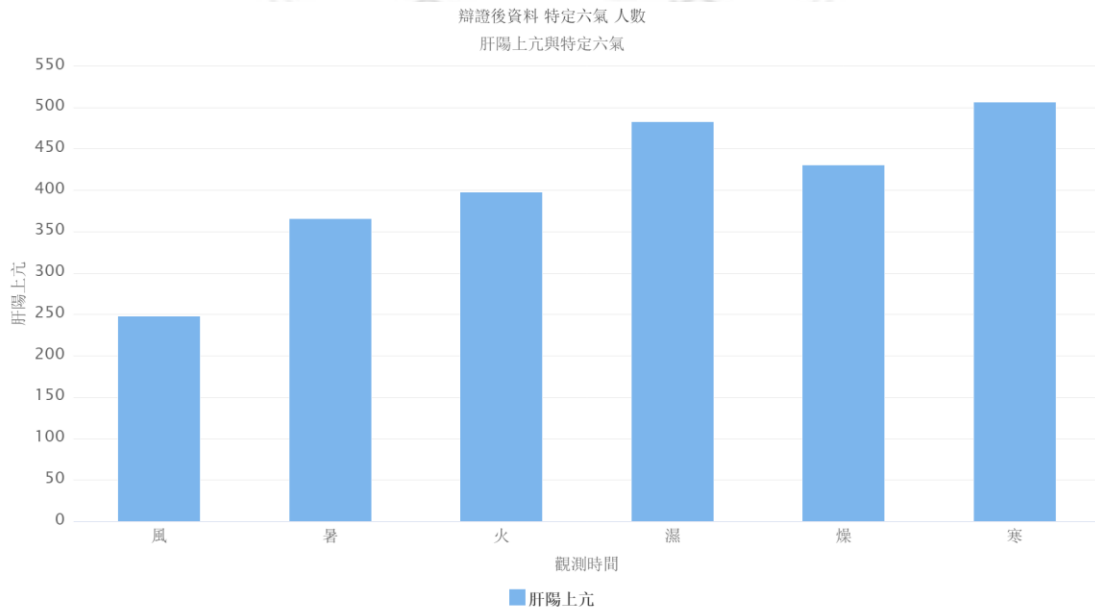


圖 81 肝陽上亢特定六氣就診人數長條圖

## 4.2. RapidMiner 資料探勘分析

經由上述敘述性統計分析過後，本節針對肝火上炎證候進行資料探勘之分析，採用 RapidMiner 資料探勘軟體中的決策樹分析，先從中醫證候與氣象因子資料方體中切出時間(每年各月)、肝火上炎與證候前一個月之各種氣象因子資料，再將資料匯入 RapidMiner 資料探勘軟體。

透過改變輸入變數、決策樹準則與訓練集比例，並進行不剪枝、事前剪枝與事後剪枝，找出最適合輸入變數為數值與類別之模型，可供後續分析不同證候時使用，如(圖 82) 所示。

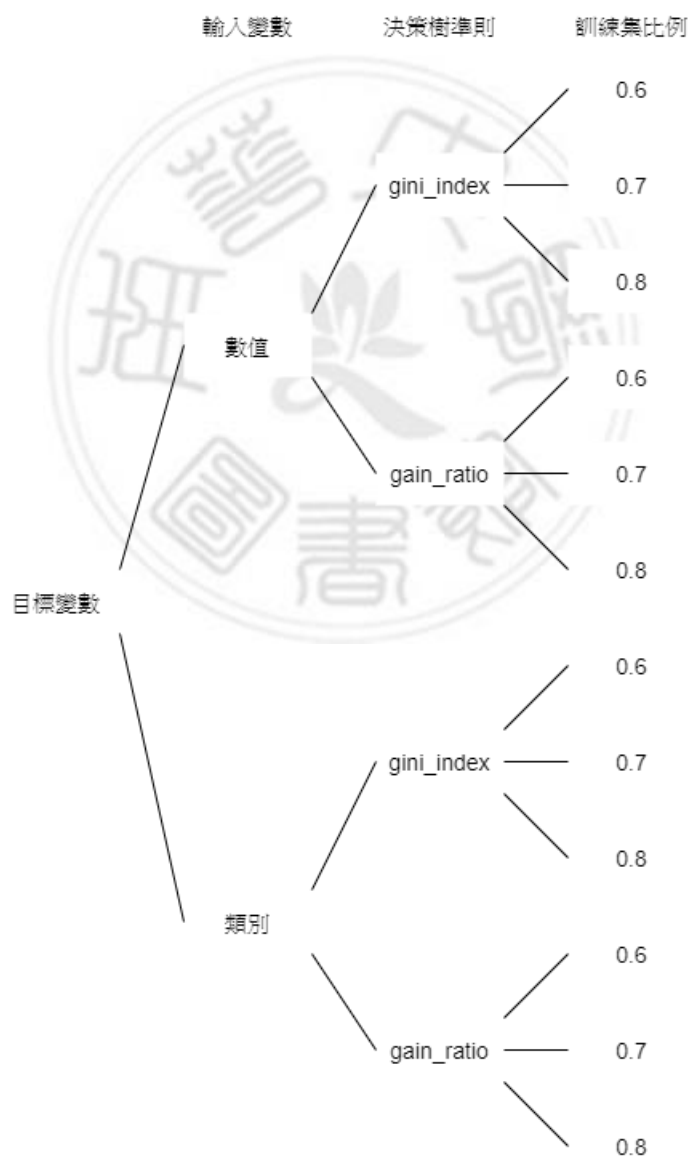


圖 82 資料探勘分析圖—以肝火上炎為例

(圖 83)為本研究 RapidMiner 資料探勘軟體之決策樹分析流程，先將目標變數設定為肝火上炎後，藉由 Multiply 將設定完之資料，分別交由三個不同 Split Data 切割成 0.6、0.7 與 0.8 三種不同比例，經由決策樹之 gini\_index 準則分析，得出三種不同的準確率，而 gain\_ratio 準則分析僅需要將決策樹內的準則更改為 gain\_ratio 準則即可。

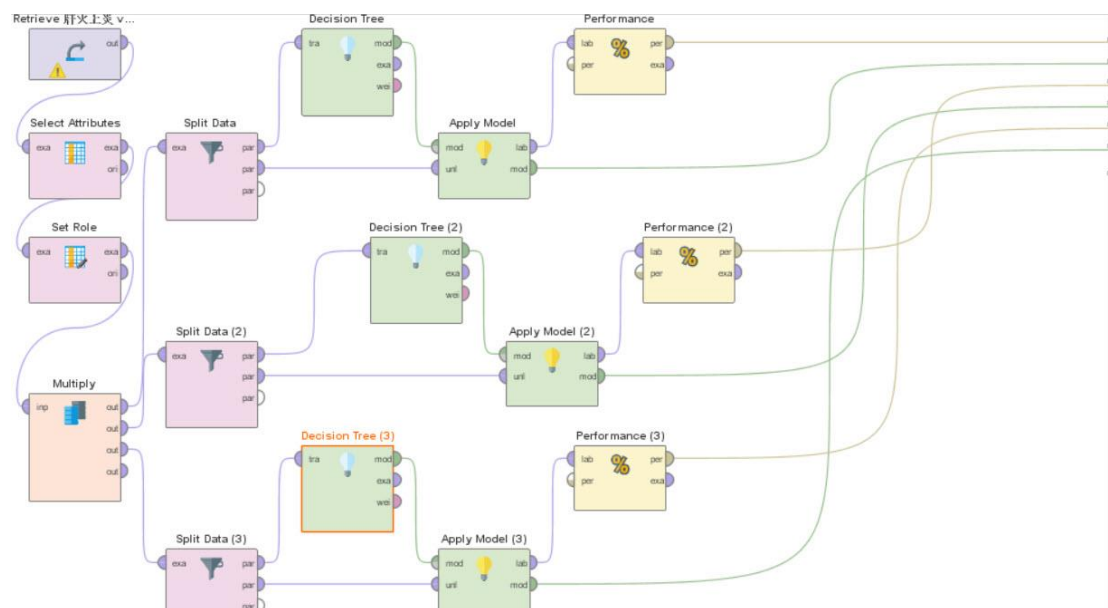


圖 83 RapidMiner 資料探勘軟體之決策樹分析流程

將事前剪枝之設定完成，將 maximal depth 設為 10，代表樹狀圖僅會有十層，minimal leaf size 設為 2，代表樹狀圖葉節點之案例最低為兩個，minimal size for split 設為 4，代表樹狀圖節點之案例最低為四個，其餘維持預設。

#### 4.2.1. 數值型態之輸入變數

首先以數值型態之輸入變數進行決策樹分析，分別採用以下三點，找出準確率最高的模型：

1. 兩種不同決策樹準則分別為 gini\_index 準則與 gain\_ratio 準則。
2. 訓練資料集比例分別為 0.6、0.7、0.8。
3. 剪枝方式分別為不剪枝、事前剪枝與事後剪枝。

(圖 84-86)為不剪枝且使用 gini\_index 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 66.67%，為不剪枝之 gini\_index 準則中最高。

accuracy: 42.11%

	true 低	true 中	true 高
pred. 低	2	3	0
pred. 中	2	1	2
pred. 高	0	4	5

圖 84 gini\_index 與訓練資料集為 0.6 之決策樹不剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	1	2	0
pred. 中	2	3	0
pred. 高	0	1	5

圖 85 gini\_index 與訓練資料集為 0.7 之決策樹不剪枝分析結果

accuracy: 66.67%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	0	1	0
pred. 高	0	3	3

圖 86 gini\_index 與訓練資料集為 0.8 之決策樹不剪枝分析結果

(圖 87-89)為事前剪枝且使用 gini\_index 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 66.67%，為 gini\_index 準則中最高。

accuracy: 47.37%

	true 低	true 中	true 高
pred. 低	2	2	0
pred. 中	2	2	2
pred. 高	0	4	5

圖 87 gini\_index 與訓練資料集為 0.6 之決策樹事前剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	1	0	0
pred. 中	2	3	0
pred. 高	0	3	5

圖 88 gini\_index 與訓練資料集為 0.7 之決策樹事前剪枝分析結果

accuracy: 66.67%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	0	1	0
pred. 高	0	3	3

圖 89 gini\_index 與訓練資料集為 0.8 之決策樹事前剪枝分析結果



(圖 90-92)為事後剪枝且使用 gini\_index 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 66.67%，為 gini\_index 準則中最高，與不剪枝一致。

accuracy: 42.11%

	true 低	true 中	true 高
pred. 低	2	3	0
pred. 中	2	1	2
pred. 高	0	4	5

圖 90 gini\_index 與訓練資料集為 0.6 之決策樹事後剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	1	2	0
pred. 中	2	3	0
pred. 高	0	1	5

圖 91 gini\_index 與訓練資料集為 0.7 之決策樹事後剪枝分析結果

accuracy: 66.67%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	0	1	0
pred. 高	0	3	3

圖 92 gini\_index 與訓練資料集為 0.8 之決策樹事後剪枝分析結果

(圖 93-95)為不剪枝且使用 gain\_ratio 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 88.89%，為 gain\_ratio 準則中最高。

accuracy: 57.89%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	2	4	2
pred. 高	0	4	5

圖 93 gain\_ratio 與訓練資料集為 0.6 之決策樹不剪枝分析結果

accuracy: 50.00%

	true 低	true 中	true 高
pred. 低	1	0	0
pred. 中	2	2	1
pred. 高	0	4	4

圖 94 gain\_ratio 與訓練資料集為 0.7 之決策樹不剪枝分析結果

accuracy: 88.89%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	0	4	1
pred. 高	0	0	2

圖 95 gain\_ratio 與訓練資料集為 0.8 之決策樹不剪枝分析結果

(圖 96-98)為事前剪枝且使用gain\_ratio準則與訓練資料集比例分別為0.6、0.7、0.8之準確度分析，由此三種分析結果可看出，在訓練資料集比例為0.6的情況下，模型準確度達到57.89%，為gain\_ratio準則中最高。

accuracy: 57.89%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	2	8	6
pred. 高	0	0	1

圖 96 gain\_ratio 與訓練資料集為 0.6 之決策樹事前剪枝分析結果

accuracy: 50.00%

	true 低	true 中	true 高
pred. 低	1	0	0
pred. 中	2	3	2
pred. 高	0	3	3

圖 97 gain\_ratio 與訓練資料集為 0.7 之決策樹事前剪枝分析結果

accuracy: 55.56%

	true 低	true 中	true 高
pred. 低	1	0	0
pred. 中	1	4	3
pred. 高	0	0	0

圖 98 gain\_ratio 與訓練資料集為 0.8 之決策樹事前剪枝分析結果

(圖 99-101)為事後剪枝且使用 gain\_ratio 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 88.89%，為 gain\_ratio 準則中最高，與不剪枝一致。

accuracy: 57.89%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	2	4	2
pred. 高	0	4	5

圖 99 gain\_ratio 與訓練資料集為 0.6 之決策樹事後剪枝分析結果

accuracy: 50.00%

	true 低	true 中	true 高
pred. 低	1	0	0
pred. 中	2	2	1
pred. 高	0	4	4

圖 100 gain\_ratio 與訓練資料集為 0.7 之決策樹事後剪枝分析結果

accuracy: 88.89%

	true 低	true 中	true 高
pred. 低	2	0	0
pred. 中	0	4	1
pred. 高	0	0	2

圖 101 gain\_ratio 與訓練資料集為 0.8 之決策樹事後剪枝分析結果

#### 4.2.2. 類別型態之輸入變數

以類別型態之輸入變數進行決策樹分析，分別採用以下三點，找出準確率最高的模型：

1. 兩種不同決策樹準則分別為 gini\_index 準則與 gain\_ratio 準則。
2. 訓練資料集比例分別為 0.6、0.7、0.8。
3. 剪枝方式分別為不剪枝、事前剪枝與事後剪枝。

(圖 102-104)為不剪枝且使用 gini\_index 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 66.67%，為 gini\_index 準則中最高。

accuracy: 47.37%

	true 低	true 中	true 高
pred. 低	3	2	2
pred. 中	1	3	2
pred. 高	0	3	3

圖 102 gini\_index 與訓練資料集為 0.6 之決策樹不剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	3	0	1
pred. 中	0	3	1
pred. 高	0	3	3

圖 103 gini\_index 與訓練資料集為 0.7 之決策樹不剪枝分析結果

accuracy: 66.67%

	true 低	true 中	true 高
pred. 低	2	1	0
pred. 中	0	2	1
pred. 高	0	1	2

圖 104 gini\_index 與訓練資料集為 0.8 之決策樹不剪枝分析結果

(圖 105-107)為事前剪枝且使用 gini\_index 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.7 情況下，模型準確度達到 85.71%，為 gini\_index 準則中最高。

accuracy: 52.63%

	true 低	true 中	true 高
pred. 低	4	2	1
pred. 中	0	4	4
pred. 高	0	2	2

圖 105 gini\_index 與訓練資料集為 0.6 之決策樹事前剪枝分析結果

accuracy: 85.71%

	true 低	true 中	true 高
pred. 低	3	1	0
pred. 中	0	4	0
pred. 高	0	1	5

圖 106 gini\_index 與訓練資料集為 0.7 之決策樹事前剪枝分析結果

accuracy: 66.67%

	true 低	true 中	true 高
pred. 低	2	1	0
pred. 中	0	2	1
pred. 高	0	1	2

圖 107 gini\_index 與訓練資料集為 0.8 之決策樹事前剪枝分析結果

(圖 108-110)為事後剪枝且使用 gini\_index 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.8 的情況下，模型準確度達到 66.67%，為 gini\_index 準則中最高，與不剪枝一致。

accuracy: 47.37%

	true 低	true 中	true 高
pred. 低	3	2	2
pred. 中	1	3	2
pred. 高	0	3	3

圖 108 gini\_index 與訓練資料集為 0.6 之決策樹事後剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	3	0	1
pred. 中	0	3	1
pred. 高	0	3	3

圖 109 gini\_index 與訓練資料集為 0.7 之決策樹事後剪枝分析結果

accuracy: 66.67%

	true 低	true 中	true 高
pred. 低	2	1	0
pred. 中	0	2	1
pred. 高	0	1	2

圖 110 gini\_index 與訓練資料集為 0.8 之決策樹事後剪枝分析結果

(圖 111-113)為使用 gain\_ratio 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.7 的情況下，模型準確度達到 64.29%，為 gain\_ratio 準則中最高。

accuracy: 63.16%

	true 低	true 中	true 高
pred. 低	4	2	2
pred. 中	0	3	0
pred. 高	0	3	5

圖 111 gain\_ratio 與訓練資料集為 0.6 之決策樹不剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	3	1	0
pred. 中	0	3	2
pred. 高	0	2	3

圖 112 gain\_ratio 與訓練資料集為 0.7 之決策樹不剪枝分析結果

accuracy: 55.56%

	true 低	true 中	true 高
pred. 低	2	2	0
pred. 中	0	1	1
pred. 高	0	1	2

圖 113 gain\_ratio 與訓練資料集為 0.8 之決策樹不剪枝分析結果



(圖 114-116)為事前剪枝且使用 gain\_ratio 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.7 的情況下，模型準確度達到 71.43%，為 gain\_ratio 準則中最高。

accuracy: 52.63%

	true 低	true 中	true 高
pred. 低	4	2	1
pred. 中	0	4	4
pred. 高	0	2	2

圖 114 gain\_ratio 與訓練資料集為 0.6 之決策樹事前剪枝分析結果

accuracy: 71.43%

	true 低	true 中	true 高
pred. 低	3	1	0
pred. 中	0	4	2
pred. 高	0	1	3

圖 115 gain\_ratio 與訓練資料集為 0.7 之決策樹事前剪枝分析結果

accuracy: 55.56%

	true 低	true 中	true 高
pred. 低	2	2	0
pred. 中	0	1	1
pred. 高	0	1	2

圖 116 gain\_ratio 與訓練資料集為 0.8 之決策樹事前剪枝分析結果

(圖 117-119)為事後剪枝且使用 gain\_ratio 準則與訓練資料集比例分別為 0.6、0.7、0.8 之準確度分析，由此三種分析結果可看出，在訓練資料集比例為 0.7 的情況下，模型準確度達到 64.29%，為 gain\_ratio 準則中最高，與不剪枝一致。

accuracy: 63.16%

	true 低	true 中	true 高
pred. 低	4	2	2
pred. 中	0	3	0
pred. 高	0	3	5

圖 117 gain\_ratio 與訓練資料集為 0.6 之決策樹事後剪枝分析結果

accuracy: 64.29%

	true 低	true 中	true 高
pred. 低	3	1	0
pred. 中	0	3	2
pred. 高	0	2	3

圖 118 gain\_ratio 與訓練資料集為 0.7 之決策樹事後剪枝分析結果

accuracy: 55.56%

	true 低	true 中	true 高
pred. 低	2	2	0
pred. 中	0	1	1
pred. 高	0	1	2

圖 119 gain\_ratio 與訓練資料集為 0.8 之決策樹事前剪枝分析結果

### 4.3. 分析結果討論

#### 4.3.1. 敘述性統計結果討論

經由上述六種時間不同維度切割方式，分別以月份、季節與六氣之統計表進行討論，如(表六)所示。

觀察證候雖然在月份維度的粗細不同下，產生的觀察情況有些微不同，除心脈痺阻(氣滯)外，其餘觀察證候皆在1月與2月偏低，而在10月、11月與12月則偏高。

腎不納氣在每年各月與其他觀察證候出現不同狀況，而在特定月份則相似，可能為特定月份是將2010年至2013年按照不同月份總和，導致在每年各月出現的特徵消失。

心脈痺阻(氣滯)雖在每年各月沒出現特別偏低的狀況，但在維度粗細的調整後，在特定月份出現2月偏低，此為維度上卷之好處，可透過上卷發現不同規律。

心脈痺阻(氣滯)在8月偏高，根據陳宗仁(2019)在其論文整理出不同西醫之解讀，發現高溫會提高冠心病風險，從本研究先前整理之氣象因子資料中，發現7-8月的氣溫偏高，故此觀察情況為合理。

表六 不同維度切割之敘述性統計表(月份)

維度切割	證候	觀察情況
每年各月	心陰虛	每年1-2月皆偏低， 且2010、2012年10月與2011、2013年11月偏高
	腎不納氣	2011、2013年1-2月偏低，且2010、2013年4月偏高
	心脈痺阻(寒凝)	每年1-2月皆偏低，且在2010年10月偏高
	心脈痺阻(氣滯)	除2012年外，其餘年8-9月皆偏高
	肝火上炎	每年1-2月皆偏低，且在每年11-12月皆偏高
	肝陽上亢	每年1-2月皆偏低，且在每年11-12月皆偏高
特定月份	心陽虛	2月偏低，而在11月偏高
	腎不納氣	1月偏低，而在11月偏高
	心脈痺阻(寒凝)	1月偏低，而在11月偏高
	心脈痺阻(氣滯)	2月偏低，而在8月偏高
	肝火上炎	2月偏低，而在12月偏高
	肝陽上亢	2月偏低，而在12月偏高

如(表七)所示，觀察證候雖然在季節維度的粗細不同下，產生的觀察情況有些微不同，除心脈痺阻(氣滯)與腎不納氣外，其餘觀察證候皆在春季偏低，而在冬季則偏高。

腎不納氣雖在每年各季沒出現特別偏低的狀況，且原先在不同年夏季與冬季有偏高，但在維度粗細的調整後，在特定季節出現夏季偏高而春季則偏低。

心脈痺阻(氣滯)在每年各季出現不同年秋季與冬季有偏高，在維度粗細的調整後，而在特定季節出現秋季偏高，可能為特定季節是將2010年至2013年按照不同季節總和，導致其在每年各季出現的特徵消失。

根據在月份討論時提到陳宗仁(2019)整理出不同西醫之解讀，發現高溫會提高冠心病風險，從氣象因子資料中發現7-8月的氣溫偏高，秋季在本研究中為7-9月，故此心脈痺阻(氣滯)觀察情況為合理。

表七 不同維度切割之敘述性統計表(季節)

維度切割	證候	觀察情況
每年各季	心陰虛	每年冬季偏高，而每年春季則偏低
	腎不納氣	2010、2013年夏季偏高，而2011、2012年冬季偏高
	心脈痺阻(寒凝)	除2011年外，其餘冬季皆偏高，而每年春季皆偏低
	心脈痺阻(氣滯)	2010-2011年秋季與2012-2013年冬季皆偏高 而每年春季則皆偏低
	肝火上炎	每年冬季偏高，而每年春季則偏低
	肝陽上亢	除2011年外，其餘年冬季偏高，而每年春季則偏低
特定季節	腎不納氣	夏季偏高，而春季則偏低
	暑襲衛表	冬季偏高，而春季則偏低
	心脈痺阻(寒凝)	冬季偏高，而春季則偏低
	心脈痺阻(氣滯)	秋季偏高，而春季則偏低
	肝火上炎	冬季偏高，而春季則偏低
	肝陽上亢	冬季偏高，而春季則偏低

如(表 八)所示，觀察證候雖然在六氣維度的粗細不同下，產生的觀察情況有些微不同，除心脈痺阻(氣滯)外，其餘觀察證候皆在風證偏低，而各自在不同六氣時偏高。

心脈痺阻(氣滯) 在每年六氣出現不同年濕證與火證有偏高，在維度粗細的調整後，而在特定六氣出現火證偏高，可能為特定六氣是將2010年至2013年按照不同六氣總和，導致在每年六氣出現的特徵消失。

根據在月份討論時提到陳宗仁(2019)整理出不同西醫之解讀，發現高溫會提高冠心病風險，從氣象因子資料中發現5-7月的氣溫偏高，火證在本研究中為5/21-7/22，故心脈痺阻(氣滯)觀察情況為合理。

表 八 不同維度切割之敘述性統計表(六氣)

維度切割	證候	觀察情況
每年六氣	腎氣不固	每年風證時偏少 而在2010-2011年寒證與2012年燥證時偏多
	寒襲衛表	每年風證時偏少 而在2010-2011年寒證與2012年燥證時偏多
	心脈痺阻(寒凝)	每年風證時偏少 而在2010年燥、2011年濕證與2012年火證時偏多
	心脈痺阻(氣滯)	2010年寒證與2011-2012年風證時偏少 2010-2011年濕證與2012年火證時偏多
	肝火上炎	每年風證時偏少，而在每年寒證時偏多
	肝陽上亢	每年風證時偏少 2010-2011年寒證與2012年濕證時偏多
特定六氣	心陰虛	風證時偏少，而在燥證時偏多
	寒襲衛表	風證時偏少，而在寒證時偏多
	心脈痺阻(寒凝)	風證時偏少，而在寒證時偏多
	心脈痺阻(氣滯)	風證時偏少，而在火證時偏多
	肝火上炎	風證時偏少，而在寒證時偏多
	肝陽上亢	風證時偏少，而在寒證時偏多

從不同時間維度切割之敘述性統計可發現在維度上卷後，雖然會使某些特徵消失或出現，但是以整體觀察來說，會使結果變為更加易懂與明確。

以肝火上炎為例，闡述在六種不同時間維度下會出現何種現象，如(表 九)所示，原先每年各月 1 月與 11 月有偏低或偏高之狀況，進行維度上卷後，在特定月份則消失，促使觀察狀況變為易懂且簡略，而季節與六氣則無消失特徵。

陳宗仁(2019)整理出肝火上炎在中醫發病原因與症狀，肝火上炎分類是一種火證，因環境溫度過熱，導致心煩氣躁所致，其標準症狀有情緒急躁、易怒、頭痛、頭暈等。

由(表 九)可看出肝火上炎會在 2 月、春季(1-3 月)與風證(1/20-3/19)時偏低，根據上述肝火上炎發病原因，從本研究之氣象因子資料中，發現 1-3 月之氣溫為偏低，故較不容易產生肝火上炎，此觀察狀況為合理。

由(表 九)可看出肝火上炎會在 12 月、冬季(10-12 月)與寒證(11/22-1/19)時偏高，據上述肝火上炎發病原因，觀察 10-12 月之氣溫發現其為偏低，並不符合發病原因，但從肝火上炎症狀來看，頭痛與頭暈好發於溫差過大(陳木森，2020)，觀察 10-12 月之溫差發現其為偏高，故以症狀來看，此觀察狀況為合理。

表 九 肝火上炎之不同時間維度統計表

證候	時間維度	觀察情況
肝火上炎	每年各月	每年 1-2 月皆偏低，且在每年 11-12 月皆偏高
	特定月份	2 月偏低，而在 12 月偏高
	每年各季	每年春季則偏低，而每年冬季偏高
	特定季節	春季則偏低，而冬季偏高
	每年六氣	每年風證時偏少，而在每年寒證時偏多
	特定六氣	風證時偏少，而在寒證時偏多

### 4.3.2. RapidMiner 資料探勘分析結果討論

經由上述分析，結果如(表 十)，不論輸入變數為何，不剪枝與事後剪枝之所有結果相同，故僅以事後剪枝與事前剪枝進行討論，以下為準確度最高之四種模型。

當輸入變數為數值型態時：

1. 使用 gain\_ratio 準則且事後剪枝，訓練資料集比例為 0.8 之模型。
2. 使用 gini\_index 準則且事前剪枝，訓練資料集比例為 0.8 之模型。

當輸入變數為類別型態時：

3. 使用 gini\_index 準則且事後剪枝，訓練資料集比例為 0.8 之模型。
4. 使用 gini\_index 準則且事前剪枝，訓練資料集比例為 0.7 之模型。

表 十 不同輸入變數型態與剪枝之決策樹分析結果

輸入變數型態	決策樹準則	訓練資料集 比例	不剪枝 準確度	事前剪枝 準確度	事後剪枝 準確度
數值	gini_index	0.6	42.11%	47.73%	42.11%
		0.7	64.29%	64.29%	64.29%
		0.8	66.67%	66.67%	66.67%
	gain_ratio	0.6	57.89%	57.89%	57.89%
		0.7	50.00%	50.00%	50.00%
		0.8	88.89%	55.56%	88.89%
類別	gini_index	0.6	47.37%	52.63%	43.73%
		0.7	64.29%	85.71%	64.29%
		0.8	66.67%	66.67%	66.67%
	gain_ratio	0.6	63.16%	52.63%	63.16%
		0.7	64.29%	71.43%	64.29%
		0.8	55.56%	55.56%	55.56%

### 4.3.2.1. 數值型態之輸入變數模型分析

本節將針對下列兩種數值型態之輸入變數模型進行討論：

1. 使用 gain\_ratio 準則且事後剪枝，訓練資料集比例為 0.8 之模型(後續簡稱為數值事後 gain\_ratio 模型)。
2. 使用 gini\_index 準則且事前剪枝，訓練資料集比例為 0.8 之模型(後續簡稱為數值事前 gini\_index 模型)。

(圖 120)為數值事後 gain\_ratio 模型之決策樹樹狀圖。



圖 120 數值事後 gain\_ratio 模型之決策樹樹狀圖



(表 十一)為數值事後 gain\_ratio 模型之決策樹規則描述，本研究觀察到測站平均氣壓、平均氣溫、最低氣溫、測站最高氣壓、懸浮微粒 PM10、測站最低氣壓與累積降水日數為有相關之變數，並分別針對低、中、高之規則進行分析與討論。

表 十一 數值事後 gain\_ratio 模型之決策樹規則描述

規則描述	肝火上炎類別
平均氣壓>1015.15	低
平均氣壓≤1015.15 & 平均氣溫≤16.75	低
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫≤6.95	低
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓 >1019.15	低
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓≤1019.15 & 懸浮微粒 PM10>33.50 & 累積降水量>375.85 & 平均氣壓>1003.7	低
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓 >1019.55	中
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓≤1019.15 & 懸浮微粒 PM10≤33.50	中
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓≤1019.15 & 懸浮微粒 PM10>33.50 & 累積降水量≤ 375.85 & 最低氣壓>992.35 & 累積降水日數>10.5	中
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓≤1019.15 & 懸浮微粒 PM10>33.50 & 累積降水量>375.85 & 平均氣壓≤ 1003.7	高
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓≤1019.15 & 懸浮微粒 PM10>33.50 & 累積降水量≤ 375.85 & 最低氣壓≤992.35	高
平均氣壓≤1015.15 & 平均氣溫>16.75 & 最低氣溫>6.95 & 最高氣壓≤1019.55 & 最高氣壓≤1019.15 & 懸浮微粒 PM10>33.50 & 累積降水量≤ 375.85 & 最低氣壓>992.35 & 累積降水日數≤10.5	高

數值事後 gain\_ratio 模型之主要節點為測站平均氣壓，以下為不同肝火上炎類別之規則：

1. 類別為低之規則有五種，影響分枝之結點分別為平均氣壓、平均氣溫與最低氣溫。
2. 類別為中之規則有三種，影響分枝之結點分別為最高氣壓、懸浮微粒 PM10 與累積降水日數。
3. 類別為高之規則有三種，影響分枝之結點分別為平均氣壓、最低氣壓與累積降水日數。

從上述可發現，除平均氣壓為影響全部類別之因子外，其餘之平均氣溫、最低氣溫、最高氣壓、最低氣壓、懸浮微粒 PM10 與累積降水量，這七種氣象因子分別影響不同類別的肝火上炎，如(表 十二)所示。

表 十二 數值事後 gain\_ratio 模型之氣象因子影響肝火上炎類別表

肝火上炎類別	影響之氣象因子
低	平均氣壓、平均氣溫與最低氣溫
中	最高氣壓、懸浮微粒 PM10 與累積降水日數
高	平均氣壓、最低氣壓與累積降水日數

接著針對數值事前 gini\_index 模型之樹狀圖進行觀察，由(圖 121)可看出，當測站平均氣壓>1015.15 時，肝火上炎類別會為低，代表當平均氣壓降低時，肝火上炎類別會提升，呈現負相關。

數值事前 gini\_index 模型之決策樹樹狀圖僅有兩層，經由觀察後，發現是因為 minimal leaf size 設為 2，許多僅有 1 之葉節點皆被剪枝，導致樹狀圖僅有兩層。



圖 121 數值事前 gini\_index 模型之決策樹樹狀圖

(表 十三)為數值事前 gini\_index 模型之決策樹規則描述，本研究觀察到僅有測站平均氣壓為有相關之變數，在平均氣壓降低時，肝火上炎類別會上升。

表 十三 數值事前 gini\_index 模型之決策樹規則描述

規則描述	肝火上炎類別
平均氣壓>1015.15	低
平均氣壓≤1015.15 & 平均氣壓>1003.05	中
平均氣壓≤1015.15 & 平均氣壓≤1003.05	高

#### 4.3.2.2. 類別型態之輸入變數模型分析

本節將針對下列兩種類別型態之輸入變數模型進行討論：

1. 使用 gini\_index 準則且事後剪枝，訓練資料集比例為 0.8 之模型(後續簡稱為類別事後 gini\_index 模型)。
2. 使用 gini\_index 準則且事前剪枝，訓練資料集比例為 0.7 之模型(後續簡稱為類別事前 gini\_index 模型)。

針對類別事後 gini\_index 模型之規則描述進行觀察，其主要結點為最低氣溫，分別依照低、中、高之規則進行討論。

(表 十四)為肝火上炎類別低之統計表，其主要節點為最低氣溫，類別低之規則有四種，影響分枝之節點分別為平均風速、平均氣溫與溫差。

表 十四 類別事後 gini\_index 模型之規則描述(肝火上炎類別低)

規則描述	肝火上炎類別
最低氣溫=低 & 平均風速=高	低
最低氣溫=低 & 平均風速=中 & 平均相對溼度=低 & 平均氣溫=低	低
最低氣溫=低 & 平均風速=中 & 平均相對溼度=高 & 最高氣溫=中 & 累積降水時數=低	低
最低氣溫=高 & 最大日降水量=中 & 溫差=中	低

(表 十五)為肝火上炎類別中之統計表，其主要節點為最低氣溫，類別中之規則有十一種，影響分枝之節點分別為平均相對溼度、最高氣溫、細懸浮微粒 PM2.5 與最大日降水量。

表 十五 類別事後 gini\_index 模型之規則描述(肝火上炎類別中)

規則描述	肝火上炎類別
最低氣溫=低 & 平均風速=中 & 平均相對溼度=中	中
最低氣溫=低 & 平均風速=中 & 平均相對溼度=低 & 平均氣溫=中	中
最低氣溫=低 & 平均風速=中 & 平均相對溼度=高 & 最高氣溫=低	中
最低氣溫=低 & 平均風速=中 & 平均相對溼度=高 & 最高氣溫=中 & 累積降水時數=中	中
最低氣溫=中 & 最高氣溫=高	中
最低氣溫=中 & 最高氣溫=中 & 細懸浮微粒 PM2.5=高	中
最低氣溫=中 & 最高氣溫=中 & 細懸浮微粒 PM2.5=低 & 累積日照時數=中	中
最低氣溫=中 & 最高氣溫=中 & 細懸浮微粒 PM2.5=低 & 累積日照時數=低	中
最低氣溫=高 & 最大日降水量=高	中
最低氣溫=高 & 最大日降水量=低 & 平均相對溼度=中	中
最低氣溫=高 & 最大日降水量=低 & 平均相對溼度=低 & 累積日照時數=中	中

(表 十六)為肝火上炎類別高之統計表，其主要節點為最低氣溫，類別高之規則有五種，影響分枝之節點分別為最高氣溫、細懸浮微粒 PM2.5 與溫差。

表 十六 類別事後 gini\_index 模型之規則描述(肝火上炎類別高)

規則描述	肝火上炎類別
最低氣溫=中 & 最高氣溫=低	高
最低氣溫=中 & 最高氣溫=中 & 細懸浮微粒 PM2.5=中	高
最低氣溫=中 & 最高氣溫=中 & 細懸浮微粒 PM2.5=低 & 累積日照時數=高	高
最低氣溫=高 & 最大日降水量=中 & 溫差=低	高
最低氣溫=高 & 最大日降水量=低 & 平均相對溼度=低 & 累積日照時數=高	高

從(表 十七)可發現，除最低氣溫為影響全部類別之因子外，其餘之平均風速、平均氣溫、溫差、平均相對溼度、最高氣溫、細懸浮微粒 PM2.5、最大日降水量與累積降水日數，這八種氣象因子分別影響不同類別的肝火上炎。

表 十七 類別事後 gini\_index 模型之氣象因子影響肝火上炎類別表

肝火上炎類別	影響之氣象因子
低	平均風速、平均氣溫與溫差
中	平均相對溼度、最高氣溫、細懸浮微粒 PM2.5 與最大日降水量
高	最高氣溫、細懸浮微粒 PM2.5 與溫差

類別事前 gini\_index 模型之主要節點為最低氣溫，(表 十八)為不同肝火上炎類別之規則：

1. 類別為低之規則有兩種，影響分枝之節點分別為平均風速與平均相對溼度。
2. 類別為中之規則有四種，影響分枝之節點分別為平均相對溼度、最低氣壓與最高氣溫。
3. 類別為高之規則有三種，影響分枝之節點分別為溫差、最低氣壓與最高氣溫。

表 十八 類別事前 gini\_index 模型之規則描述

規則描述	肝火上炎類別
最低氣溫=低 & 平均風速=高	低
最低氣溫=低 & 平均風速=中 & 平均相對溼度=低	低
最低氣溫=高 & 最低氣壓=中	中
最低氣溫=低 & 平均風速=中 & 平均相對溼度=中	中
最低氣溫=低 & 平均風速=中 & 平均相對溼度=高	中
最低氣溫=中 & 溫差=中 & 最高氣溫=高	中
最低氣溫=中 & 溫差=低	高
最低氣溫=高 & 最低氣壓=低	高
最低氣溫=中 & 溫差=中 & 最高氣溫=中	高

從上述可發現，除最低氣溫為影響全部類別之因子外，其餘之平均風速、溫差、平均相對溼度、最低氣壓與最高氣溫，這五種氣象因子分別影響不同類別的肝火上炎，如(表 十九)所示。

表 十九 類別事前 gini\_index 模型之氣象因子影響肝火上炎類別表

肝火上炎類別	影響之氣象因子
低	平均風速與平均相對溼度
中	平均相對溼度、最低氣壓與最高氣溫
高	溫差、最低氣壓與最高氣溫

### 4.3.3. 敘述性統計與模型綜合討論

#### 4.3.3.1. 數值事後 gain\_ratio 模型

依據肝火上炎之每年各月敘述性統計結合數值事後 gain\_ratio 模型之規則描述，可得出下列結論：

肝火上炎會在 1-2 月偏低，故需在 12、1 月注意以下規則：

1. 平均氣壓 > 1015.15 hpa
2. 平均氣壓 ≤ 1015.15 hpa & 平均氣溫 ≤ 16.75 °C
3. 平均氣壓 ≤ 1015.15 hpa & 平均氣溫 > 16.75 °C & 最低氣溫 ≤ 6.95 °C
4. 平均氣壓 ≤ 1015.15 hpa & 平均氣溫 > 16.75 °C & 最低氣溫 > 6.95 °C & 最高氣壓 ≤ 1019.55 hpa & 最高氣壓 ≤ 1019.15 hpa
5. 平均氣壓 ≤ 1015.15 hpa & 平均氣溫 > 16.75 °C & 最低氣溫 > 6.95 °C & 最高氣壓 ≤ 1019.55 hpa & 最高氣壓 ≤ 1019.15 hpa & 懸浮微粒 PM10 > 33.50  $\mu\text{g}/\text{m}^3$  & 累積降水量 > 375.85 mm & 平均氣壓 > 1003.7 hpa

肝火上炎會在 11-12 月偏高，故需在 10、11 月以平均氣壓 ≤ 1015.15 hpa & 平均氣溫 > 16.75 °C & 最低氣溫 > 6.95 °C & 最高氣壓 ≤ 1019.55 hpa & 最高氣壓 ≤ 1019.15 hpa & 懸浮微粒 PM10 > 33.50 為基礎，注意以下規則

1. 累積降水量 > 375.85 mm & 平均氣壓 ≤ 1003.7 hpa
2. 累積降水量 ≤ 375.85 mm & 最低氣壓 ≤ 992.35 hpa
3. 累積降水量 ≤ 375.85 mm & 最低氣壓 > 992.35 hpa & 累積降水日數 ≤ 10.5 日

#### 4.3.3.2. 數值事前 gini\_index 模型

本研究僅觀察到測站平均氣壓為有相關之變數，在平均氣壓降低時，肝火上炎類別會上升，故不與敘述性統計進行統合。



#### 4.3.3.3. 類別事後 gini\_index 模型

依據肝火上炎之每年各月敘述性統計結合類別事後 gini\_index 模型之規則描述，可得出下列結論：

肝火上炎會在 1-2 月偏低，故需在 12、1 月注意以下規則：

1. 最低氣溫在 5.6~12.2 °C & 平均風速在 2.4~3 m/s
2. 最低氣溫在 5.6~12.2 °C & 平均風速在 1.9~2.3 m/s & 平均相對溼度在 72~76% & 平均氣溫在 14.7~19.9 °C
3. 最低氣溫在 5.6~12.2 °C & 平均風速在 1.9~2.3 m/s & 平均相對溼度在 81~84% & 最高氣溫在 29.3~33.1 °C & 累積降水時數在 0.3~43.3 小時
4. 最低氣溫在 19.2~24 °C & 最大日降水量在 51.2~113mm & 溫差在 15.5~20 °C

肝火上炎會在 11-12 月偏高，故需在 10、11 月注意以下規則

1. 最低氣溫在 12.8~18.4 °C & 最高氣溫在 23.5~28.6 °C
2. 最低氣溫在 12.8~18.4 °C & 最高氣溫在 29.3~33.1 °C & 細懸浮微粒 PM2.5 在 35~5  $\mu\text{g}/\text{m}^3$
3. 最低氣溫在 12.8~18.4 °C & 最高氣溫在 29.3~33.1 °C & 細懸浮微粒 PM2.5 在 35~5  $\mu\text{g}/\text{m}^3$  & 累積日照時數在 191.1~235.8 小時
4. 最低氣溫在 19.2~24 °C & 最大日降水量在 51.2~113mm & 溫差在 10.5~14.5 °C
5. 最低氣溫在 19.2~24 °C & 最大日降水量在 -9.8~43.7mm & 平均相對溼度在 72~76% & 累積日照時數在 191.1~235.8 小時

#### 4.3.3.4. 類別事前 gini\_index 模型

依據肝火上炎之每年各月敘述性統計結合類別事前 gini\_index 模型之規則描述，可得出下列結論：

肝火上炎會在 1-2 月偏低，故需在 12、1 月注意以下規則：

1. 最低氣溫在 5.6~12.2 °C & 平均風速在 2.4~3 m/s
2. 最低氣溫在 5.6~12.2 °C & 平均風速在 1.9~2.3 m/s & 平均相對溼度在 72~76%

肝火上炎會在 11-12 月偏高，故需在 10、11 月注意以下規則

1. 最低氣溫在 12.8~18.4 °C & 溫差在 10.5~14.5 °C
2. 最低氣溫在 19.2~24 °C & 最低氣壓在 982.2~1001.2 hpa
3. 最低氣溫在 12.8~18.4 °C & 溫差在 15.5~20 °C & 最高氣溫在 29.3~33.1 °C

#### 4.3.4. 相關論文之比較與討論

本研究因有參考陳宗仁(2019)之研究，故與其研究進行比較與討論，希冀與不同論文進行比較與討論找出差異之處，有利於後續研究者進行相關研究。

##### 4.3.4.1. 資料探勘方法比較

如(表二十)所示，兩個研究之差異除資料探勘方法不同外，對於氣象因子與證候之關係也是主要差異。

表二十 資料探勘方法之比較表

	本研究	陳宗仁之研究
採用方法	敘述性統計	Pearson 相關係數
	決策樹分析	單因子變異數分析
內容	1. 以不同時間維度進行敘述性統計找出在不同時間維度下有何差異 2. 藉由決策樹分析找出多種氣象因子與肝火上炎之關係與規則	1. 藉由 Pearson 相關係數找出單氣象因子與證候母體之線性關係 2. 藉由單因子變異數分析驗證 Pearson 相關係數找出之關係是否正確
差異	注重多氣象因子與證候之關係	注重單氣象因子與證候之關係

##### 4.3.4.2. 資料探勘結果比較

如(表二十一)所示，陳宗仁之研究結果皆為單氣象因子與證候之關係，而本研究僅有數值事前 gini\_index 模型之結果與其中一項相同，皆發現肝火上炎與平均氣壓有負相關之事實，其餘模型之結果皆為多氣象因子與證候之規則。

表二十一 不同論文之研究結果

	陳宗仁之研究
結果	心脈痺阻(寒凝)與平均風速有負相關 心脈痺阻(氣滯)與平均氣溫有正相關 肝火上炎與平均氣壓有負相關 肝火上炎與平均氣溫有正相關 肝陽上亢與平均氣溫有正相關

#### 4.3.5. 證候人數百分比公式比較

本研究證候人數百分比公式為每個月證候人數除上該年證候總人數，此公式為找出每個月特定證候在該年之人數比例，故分子與分母皆是同一特定證候，每個月證候人數會受到看診日數影響，但每年皆有固定連假與休診日，故容易產生某些月份人數偏少的問題。

本研究在完成敘述性統計與決策樹分析後發現上述問題，故提出不一樣之證候人數百分比公式，以該月特定證候除以該月全部總證候人數，找出特定證候在該月人數百分比，即可避免某些月份人數偏少的問題，希冀後續研究團隊可以此方式繼續進行。



## 第五章結論與未來發展

### 5.1. 結論

本研究利用資料倉儲之萃取、轉換與載入概念，整合中醫證候與氣象因子資料，建構中醫證候與氣象因子資料方體，再從此資料方體內，切割中醫證候與各種不同時間維度進行敘述性統計，透過資料方體之應用，了解證候在不同時間維度會出現何種特徵，同時展示資料方體在不同粗細維度會出現與消失不同特徵，最後使用 Rapid Miner 資料探勘軟體的決策樹分析法，透過使不同資料型態、準則、訓練資料集比例與剪枝方式，找出適合不同型態之輸入變數模型，結合兩種不同分析歸納出研究結果與後續改善。

#### 5.1.1. 研究結果

在敘述性統計結合不同型態之輸入變數模型中，找出在 12、1 月與 10、11 月時，平均氣壓最影響數值型態與最低氣溫最影響類別型態，與其他氣象因子在不同規則下皆會影響肝火上炎證候。

雖與陳宗仁之研究結果不相同，但透過比較與討論可了解兩者之間的差異，本研究為多氣象因子與證候之規則與關係。

#### 5.1.2. 後續改善

本研究在完成研究後，發現證候人數問題，故提出不同證候人數百分比公式，以該月特定證候除以該月全部總證候人數，找出特定證候在該月人數百分比，即可避免某些月份人數偏少的問題，希冀後續研究團隊可以此方式繼續進行。

## 5.2. 未來發展

本研究採用嘉義大林慈院中醫部的門診資料，希冀後續能取得更多不同來源的中醫證候資料，並採用年份能更多，將資料方體逐漸擴大，並在資料預處理時，能以本研究提出的不同證候人數百分比公式進行處理，希冀能找出在不同預處理下，結果會有如何改變，進一步去探討原因。

在決策樹分析部份，利用簡單明瞭之樹狀圖與規則，找出多氣象因子與證候之間的關係，希冀後續能以不同的資料探勘演算法，像是關聯法則、時間序列與羅吉斯回歸去進行不同的分析，找出更多不同的規則，驗證本研究之規則與結果是否相符或有所不同的發現，並能以本研究找出的不同型態之輸入變數模型，對其餘證候進行分析，找出更多相關規則。



## 參考文獻

### 中文文獻

1. 鄧莉雅(2015),「決策數分析與羅吉斯回歸於資料探勘的整合運用以人事資料與民眾健康影響因素之探討為例」,國立臺灣師範大學管理學院全球經營與策略研究所碩士論文。
2. 林宇恆(2016),「決策樹結合複迴歸模型預測氣溫與雨量」,國立臺灣師範大學電機工程學系碩士論文。
3. 陳有忠(2008),「決策樹運用於銀行之詐騙帳戶」,國立東華大學資訊工程學系碩士論文。
4. 許政文(2019),「以決策樹建構選股策略--以台灣上市公司為例」,國立高雄師範大學事業經營學系碩士論文。
5. 陳葦(2019),「使用過取樣與集成學習方法的糖尿病風險估測函數」,朝陽科技大學財務金融系碩士論文。
6. 顏利憲(2013),「以決策樹分析鐵路誤點原因及解決方法」,國立成功大學交通管理科學研究所碩士論文。
7. 巫美鳳(2015),「應用資料探勘技術評估術後病人疼痛緩解時間」,國立中正大學資訊管理研究所碩士論文。
8. 吳芮妍(2019),「應用機器學習分類技術分析軍人自殺風險因素」,國防大學政治作戰學院心理及社會工作學系社工碩士班碩士論文。
9. 葉子維(2018),「顧客消費行為分析及行動銀行使用預測-決策樹、隨機森林與判別分析之比較」,國立臺北大學統計學系(所)碩士論文。
10. 劉宏偉(2011),「基於中醫運氣學理論及氣象資料採擷的疫病流行規律研究」,北京中醫藥大學碩士研究生學位論文。
11. 陳宗仁(2019),「氣象因子與癌症病患中醫實證的分析」,國立中正大學資訊工程研究所碩士論文。

12. 陳世波、劉保延、王永炎、周雪忠、倪青、何麗云(2009),「症及其分類在中醫臨床評價中的作用分析」,中醫雜誌,50(3),P.5-7。
13. 黃晶一(2015),「氣象因素與成人社區獲得性肺炎中醫證型相關性研究」,廣州中醫藥大學碩士學位論文。
14. 李淑玲(2010),「資料探勘應用於點餐系統之研究-以員工餐廳為例」,國立高雄應用科技大學資訊管理學系碩士班碩士論文。
15. 簡睿志(2011),「應用資料探勘於藥物交互作用與處方分析之研究」,明志科技大學工業工程與管理研究所碩士論文。
16. 李卓翰(2003),「資料倉儲理論與實務」,學貫行銷股份有限公司。
17. 傅梓瑜(2017),住「宅用電與時間電價資訊呈現:使用資訊面板與資料倉儲技術」,國立中央大學土木工程學系碩士論文。
18. 鄭旺昌(2012),「資料倉儲產品評選研究」,國立中山大學資訊管理研究所碩士在職專班碩士論文。
19. 黃熙勛(2013),「資料倉儲建構訂單分配決策支援系統之研究-以人造皮革產業為例」,國立中山大學資訊管理學系碩士在職專班碩士論文。
20. 楊蕙菁(2004),「資訊也是商品統一超商砸四十億找出氣溫和銷售關係」,商業週刊,850。
21. 李義偉(2004),「在XML資料倉儲中實體化資料方體選取之研究」,大葉大學資訊工程學系碩士班論文。
22. 陳品辰(2009),「應用資料探勘技術進行學生缺曠課對成績影響之研究」,明新科技大學資訊管理所碩士學位論文。
23. 胡建勳、劉堯欽(2008),「資料探勘於病患就診記錄之知識管理研究-以疾病與氣候關聯性為例」,2008 知識社群與系統發展研討會。
24. 陳木森(2020),「春天早晚溫差大 民眾如有眩暈症病史應慎防復發」,NOW健康報導。



## 英文文獻

1. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996) , From data mining to knowledge discovery in databases. American Association for Artificial Intelligence, Fall, P.37-51 .
2. Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992) , Knowledge discovery in databases: An overview. AI Magazine, 13(3), P. 57-69 .
3. Gruppe, F. H., & Owrang, M. M. (1995) , Data base mining. Information Systems Management, 12(4), P. 26 .

