

國立政治大學資訊管理學系

碩士學位論文

指導教授：莊皓鈞博士、周彥君博士

零售商業分析：購物籃資料的指數隨機圖模型

Retail Business Analytics: Exponential Random Graph  
Modeling of Market Basket Data

研究生：張月馨 撰

中華民國 108 年 09 月

## Abstract

Nowadays, market basket analysis plays an important role in retail business analysis, as it allows the retailer to develop a better understanding of consumers' purchasing tendency. However, market basket analysis lacks general rules to explain the potential reasons why the products are bought together. Therefore, this research uses Exponential Random Graph Model (ERGM) to enhance the explanatory power on discovered co-purchase relationships. The ERGM is a technique for assessing interdependencies between nodes in random graphs or networks, and it enables analysts to uncover structural features in networks. With more than three million transaction records of a leading convenience store in Taiwan, our research focuses on how to model these transaction data using ERGM and combines the Dyadic Dependence Model (Hunter, Handcock, Butts, Goodreau, & Morris, 2008) to design a new analysis process. The proposed process is aimed at guiding retailers to develop better marketing strategies regarding bundle selling/co-purchase.

Keyword: Retailing, ERGM, Dyadic Dependence Model, Product Network

## 摘要

購物籃分析在當代的零售商業分析扮演重要的角色，可以幫助零售商了解消費者之購物傾向，但是購物籃分析缺乏一般化的規則解釋商品彼此併買的潛在原因，因此本研究採用指數隨機圖模型(Exponential Random Graph Modeling, ERGM)解決購物籃分析對商品連結缺乏解釋性的限制。指數隨機圖模型是用來檢測隨機圖或是網路圖模型中彼此連結關係模式的工具，對欲解釋之網路圖結構特徵提供良好的分析方法。本研究主要探討如何應用超商零售業之交易資料，設計一套以指數隨機圖模型為基礎，加入結構特徵之二元依賴模型(Hunter, Handcock, Butts, Goodreau, & Morris, 2008)之分析應用流程，幫助零售業者對行銷策略提供更好的應用方向。

**關鍵字：**零售業、指數隨機圖模型、二元依賴模型、商品網路圖

# 目錄

第一章 緒論.....	1
第二章 文獻回顧.....	3
第一節 關聯規則(Association Rule).....	3
第二節 社群網路分析(Social Network Analysis, SNA).....	6
第三章 模型介紹.....	9
第四章 ERGM 模型分析結果.....	12
第一節 資料敘述.....	12
第二節 資料分析結果.....	19
第五章 結論.....	34
參考文獻.....	36



## 表目錄

表 1 每張發票金額及購買商品件數之敘述統計.....	13
表 2 全部商品交易資料「品號-品名稱」與「群號-群名稱」統計.....	14
表 3 節點屬性變數「商品金額」之敘述統計.....	21
表 4 節點屬性變數「商品所屬品號-品名稱」之敘述統計.....	21
表 5 節點屬性變數「商品所屬群號-群名稱」之敘述統計.....	21
表 6 部分品號與基準商品勝算比比較表.....	27
表 7 前 10 項容易一起被購買之商品品號.....	31



## 圖目錄

圖 1 Apriori 找出高頻項目組示範流程圖.....	5
圖 2 部份商品交易資料.....	12
圖 3 各主商圈之交易金額.....	13
圖 4 流程架構圖.....	19
圖 5 資料轉換範例圖.....	20
圖 6 網路架構資訊.....	21
圖 7 初始網路架構模型結果.....	22
圖 8 加入 nodecov 函數之網路架構模型結果.....	23
圖 9 加入 nodefactor 函數之網路架構模型結果.....	26
圖 10 加入 nodematch 函數之網路架構模型結果.....	28
圖 11 加入 nodematch 及 diff = TRUE 函數之網路架構模型結果.....	31
圖 12 促銷活動內容.....	33
圖 13 加入 nodemix 函數之網路架構模型結果.....	33

# 第一章 緒論

## 一、研究背景

隨著資料量急速成長，資料取得成本越來越低、儲存軟、硬體設備成本下降、軟體及人員技術的進步等因素，使得大數據分析越來越普及於企業及商業市場中(Akter & Wamba, 2016)，透過大數據分析的結果可以探索資料間的關聯性，或是找出新的關聯規則創造新的商業模式，甚或是能做到精準預測，幫助企業做到更好的銷售模式。自從 1990 年代美國 Walmart 超市利用大數據關聯規則(Association Rule)找出啤酒的銷售量受到尿布影響，並且成功利用關聯規則的結果提高啤酒與尿布的銷售量，關聯規則便大量的應用在零售或電商的購物分析，希望能找出有趣且有效的關聯規則及模式，給予企業更好的決策制定參考。

關聯規則應用在零售市場中的交易資料稱為購物籃分析(Market Basket Analysis)。購物籃分析，在近幾十年被大量運用，零售業者對於消費者的購買傾向極感興趣，例如哪幾項商品會同時被購買？當知道哪幾項商品容易被購買也影響著零售業者的行銷策略，是不是採取併買行銷策略會提升銷售量？或是改變零售商店的商品陳設位置也能改變銷售量？透過購物籃分析，能夠回答購買者的消費傾向，提供零售業者更好的決策選擇與制定，進而提高銷售量。

## 二、研究動機及目的

面對市場的快速變化，零售業者希望能將大量的寶貴資料透過數據分析抓住市場機會，雖然購物籃分析在實務方面可以找到有用及有趣的資訊，在學術界上也有許多關聯規則的研究(Agrawal & Srikant, 1994; Chen, Tang, Shen, & Hu, 2005)，但卻仍有許多無法解釋的關聯規則，舉例來說，零售業者雖然

知道哪幾項商品會同時被購買，卻無法透過傳統的購物籃分析解釋是什麼原因讓它們彼此被購買。

為了解決購物籃分析的問題，本研究希望將商品交易資料從傳統的關聯規則轉為網路層面架構的商品網路分析，其中商品網路分析(Product Network Analysis)是由 Raeder and Chawla (2009)提出，其觀念是將每個商品視為一個網路節點，當商品與商品之間同時被購買就形成一條連線，也就是所謂的併買。本研究採用由 Frank and Strauss (1986)及 Wasserman and Pattison (1996)所提出的指數隨機圖模型(Exponential Random Graph Modeling)，在商品網路分析的基礎下，賦予每個商品節點屬性變數，能夠預測及解釋商品連結的機率。利用商品網路分析能夠強化傳統關聯規則下的探索性及敘述性分析結果，也解決關聯規則缺乏一般化的規則解釋商品彼此併買原因之限制，希望透過指數隨機圖模型幫助零售業者對商品彼此的連結機率有更進一步的理解。

### 三、研究問題

本論文的研究問題為：應如何實際應用指數隨機圖模型於零售流通業之併買分析？本研究將實際以國內知名超商之商品交易資料加以驗證，透過設計一套指數隨機圖模型運用於零售交易資料分析的流程，對商品交易資料併買分析提供更好的方式，以期對研究問題能有更進一步探討。

## 第二章 文獻回顧

### 第一節 關聯規則(Association Rule)

Agrawal et al. (1993)最先提出關聯規則主要目的是從交易紀錄資料庫中，分析出產品之間關聯性，從關聯規則中找出消費者之消費傾向，了解顧客潛在消費偏好及習慣，以提供有用的資訊幫助企業做行銷活動及庫存預測等作業(Kaur & Kang, 2016)。在零售市場中，關聯規則被運用在購物籃分析(Market-basket Analysis)，找出顧客購物籃中哪些商品同時被購買(Agrawal & Srikant, 1994; Chen, Tang, Shen, & Hu, 2005)，以幫助賣家找出最常購買商品的組合，做出有效的交叉銷售(Cross-selling)、搭配銷售(Bundle-selling)及市場陳設等行銷方式(Mostafa, 2015)。關聯規則同時可以應用在商品推薦方面，以線上書店為例，可以透過消費者在網站上購買書籍的消費紀錄，利用關聯規則推薦消費者更多相關書籍，以此提高網站營收(Tewari, Kumar & Barman, 2014)。

關聯規則的定義描述如下：令  $I = \{i_1, i_2, \dots, i_m\}$ ， $I$  是所有商品項目(Items)的集合， $T$  是全部交易資料的集合， $T$  為  $I$  的子集合( $T \subseteq I$ )，而一筆交易資料  $T_j$  是由一個或以上的商品項目所組成的集合，稱之為項目組 (Itemsets =  $\{item_1, item_2, \dots, item_n\}$ )，當項目組  $X$  與  $Y$  被表示成  $X \rightarrow Y$ ， $X \subseteq I$ ， $Y \subseteq I$ ， $X \cap Y = \emptyset$ ，表示包含  $X$  項目的交易也會同時包含  $Y$  項目，其中  $X$  稱為前置項目組(Antecedent)， $Y$  稱為後置項目組(Consequent)。在此關聯規則中包含三個重要的參數來決定關聯規則是否成立：

1. 支持度(Support)： $Support\{X \Rightarrow Y\}$  表示在所有的交易資料中  $X$  與  $Y$  同時出現的機率  $P(X \cap Y)$ 。其中最小支持度(Min Support)是在進行關聯規則前

會先訂一個支持度的門檻值(Threshold)，此支持度的門檻值稱為最小支持度。

2. 信賴度(Confidence)： $\text{Confidence}\{X \Rightarrow Y\}$ 表示在包含有  $X$  的交易資料中也同時包含有  $Y$  的比率值  $P(Y|X)$ 。其中最小信賴度(Min Confidence)是在進行關聯規則前會先訂一個信賴度的門檻值，此信賴度的門檻值稱為最小信賴度。
3. 增益(Lift)： $\text{Lift}\{X \Rightarrow Y\}$ 表示相比於從所有項目組中找  $Y$ ，在存在  $X$  的項目組中找到  $Y$  可能性提升的倍數  $P(Y|X) / P(Y)$ 。

關聯規則的過程可以分成兩個階段：首先找出所有支持度值高於最小支持度門檻值之高頻項目組(Frequent Items)，然後根據前階段所找出的高頻項目組，以最小信賴度門檻值及增益閾值為條件，計算出所有符合的強關聯規則。找出高頻項目組演算法中最著名的便是由 Agrawal and Srikant (1994)提出的 Apriori 演算法，由圖 1 可以看出 Apriori 演算法首先從商品項目集合搜尋出各屬性的出現次數，產生候選項目集合(Candidate Itemsets) $C_1$ ，然後從候選項目集合中找出通過最小支持度值的高頻項目集合 $F_1$ (以圖 1 之例子就是滿足大於最小支持度值 2 的項目集合)，接著將 $F_1$ 內所有項目集合經過排列組合產生 $C_2$ 集合，並以原始交易資料各項目出現次數計算項目之次數，再以相同方式找出高於最小支持度值之 $F_2$ 集合，依此計算下去，直到沒有 $C_n$ 為止。當找完所有高頻項目組，檢查若 $(k-1)$ 項目組為  $k$  項目組的子集合，則將 $(k-1)$ 項目組刪除，以圖 1 的例子來看，因為 $F_2$ 之 BC、BE、CE 為 $F_3$ 之 BCE 的子集合，故刪除 BC、BE 及 CE，最後得出高頻項目組合 AC 及 BCE。

Apriori 演算法以疊代(Iteration)的方式找出高頻項目組之後，最後根據滿足給定的最小信賴度及增益值找出強關聯規則。以圖 1 為例，假定最小信賴度值為

75%，增益值大於 1 才有行銷價值，當找出高頻組合 AC 及 BCE 之後，計算  $\text{Confidence}\{A \Rightarrow C\} = (1/2)/(1/2) = 1$ ;  $\text{Lift}\{A \Rightarrow C\} = 1/(3/4) = 4/3$ ，AC 項目集合滿足給定的最小信賴度及增益值，也就是最小信賴度值大於 75% 且增益值大於 1，而  $\text{Confidence}\{B\&C \Rightarrow E\} = (1/2)/(1/2) = 1$ ;  $\text{Lift}\{B\&C \Rightarrow E\} = 1/(3/4) = 4/3$ ，BCE 項目集合也符合滿足給定的最小信賴度及增益值，故項目集合 AC 及 BCE 為圖 1 例子交易資料中的強關聯規則。

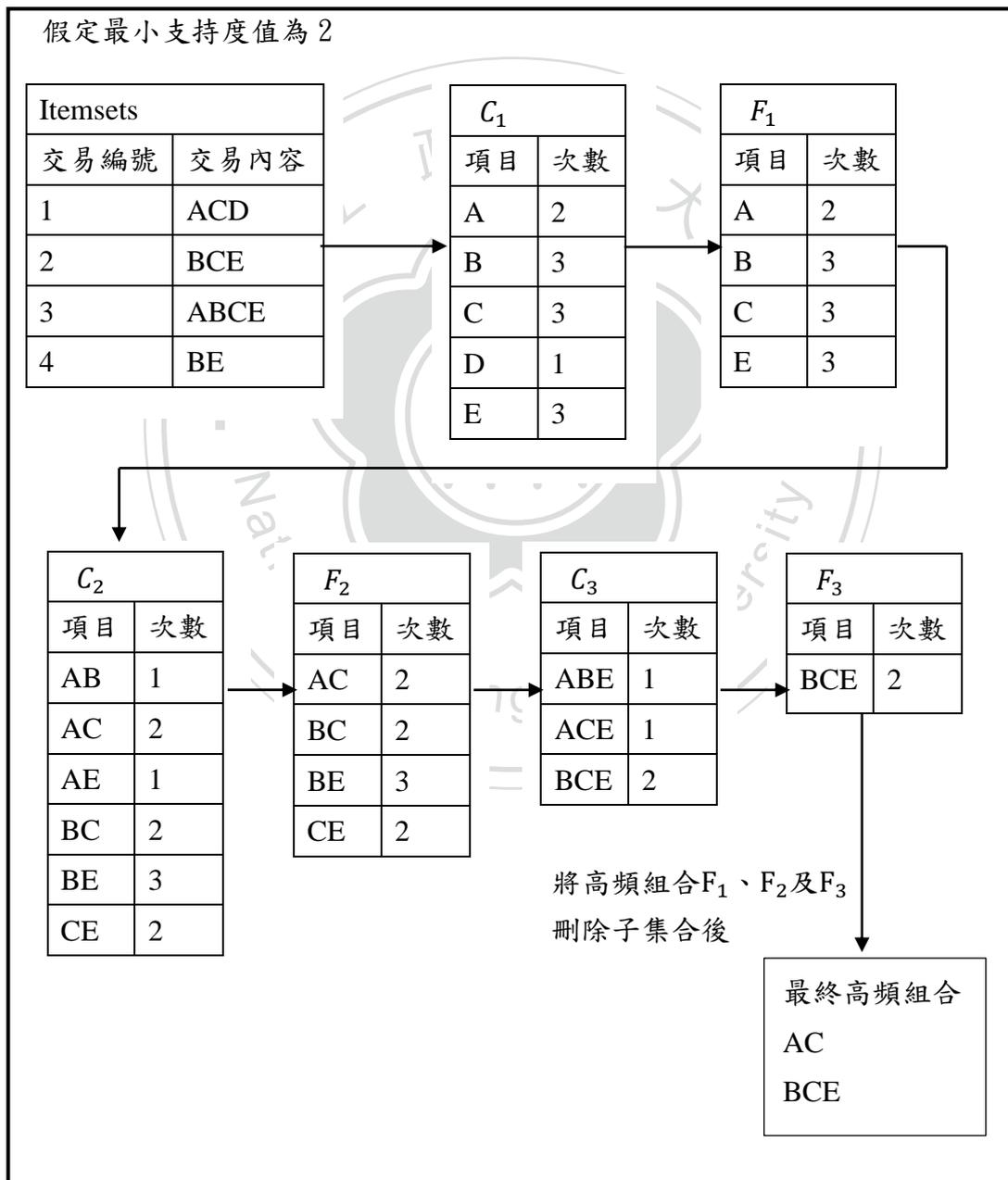


圖 1 Apriori 找出高頻項目組示範流程圖

## 第二節 社群網路分析(Social Network Analysis, SNA)

Mitchell (1969)提出社群網路分析主要是探討社群節點(Social Actor)彼此之間連結的架構關係,可以將此種關係視為一張社群網路圖(Social Network Graph),社群節點涵蓋的範圍可能是人、組織機構、由人類活動產生的商品、網路等等(Wellman and Berkowitz, 1988; Scott, 1991; Wasserman and Faust, 1994)。SNA 對於社會結構調查是一項重要的分析策略(Otte & Rousseau, 2002),可以用來查明社會結構下潛在的節點關聯模式(Qi, Fuller, Wu, Wu, & Zhang, 2012)。早期的 SNA 研究主題如醫學與農業創新(Coleman, Menzel, & Katz, 1996; Valente, 1995),在那之後不同學科研究者不論是生物學、地質學、經濟學等都對社群網路分析的主題進行研究。

SNA 不只可以透過網路圖來表達,還可以透過代表性的指標作為衡量標準,其中以網路密度(Network Density)及中心性(Centrality)兩種指標最常使用。Wasserman and Faust (1994)提出網路密度代表社群節點之間互動的連結程度,用以衡量社群網路裡節點之間連結關係的緊密程度。中心性則是用來衡量節點在網路圖中的重要程度,當節點越靠近網路圖的中心代表其重要性越高。Freeman (1979)對中心性分為三種指標:

1. 程度中心性(Degree Centrality): 當某個節點與其他節點有越多的連結,其程度中心性越高。
2. 接近中心性(Closeness Centrality): 當某個節點到網路中其他節點的最短距離總和最小,其接近中心性越高。
3. 中介中心性(Betweenness Centrality): 當某個節點在網路中任意兩個節點之間最短路徑的橋樑次數最多,也就是充當中介的次數最高,其中介中心性越高。

早期的 SNA 主要是透過問卷、訪問等投入大量人力的方式取得資料來研究特定社群網路現象，到了 1990 年代研究者開始利用科技來取得大量資料來探討更複雜的網路，以此能夠對更多領域做 SNA，包含大眾交通網路、金融客戶網路等。一直到最近幾年網際網路的便利性及社群網路平台的興起，如 Facebook、LinkedIn、Line 等，讓社群資料取得更加容易(Bonchi, Castillo, Gionis, & Jaimes, 2011)，SNA 的進步也使得計算社會科學(Computational Social Science)出現，透過電腦運算方法以模擬、建模等方式來觀察社群網路之間彼此互動及關係(Lazer et al., 2009; Watts, 2004)。

SNA 可以應用在許多領域，在網路拍賣詐欺偵測上，SNA 可以將每筆網路拍賣交易資料中賣家對買家的交易關係做成網路圖，利用中心性、資料探勘分類等方法察覺出潛在的詐欺用戶，以此降低詐欺造成的財物損失(Chiu, Ku, Lie, & Chen, 2011)；在社群網路平台方面以 Facebook 為例，可以透過用戶在 Facebook 上用戶與用戶之間的按讚、分享、評論等資訊傳播，將此過程轉換為一個巨大的網路圖，以此找出網路圖中具影響力的節點(Jin, 2013)，假如要進行口碑式行銷或是病毒式行銷都可以以他們為起點進行資訊散播，達成最大化影響力；在企業組織方面，Balkundi and Kilduff (2006)提到 SNA 可以讓組織在遇到財務問題時，獲得重要的見解並且找出生存策略。

由 Raeder and Chawla (2009)提出當 SNA 的社群節點為交易商品時，就能對商品之間的連結關係做出商品網路分析(Product Network Analysis)。相較於傳統關聯規則中找出的併買事實分析結果，商品網路分析能夠對交易資料作出更多的分析應用，例如：在產品推薦系統優化方面，能夠解決協同過濾(Collaborative Filtering)所收集到之交易或是回饋資料稀疏性問題(Sparsity Problem)，透過商品網路分析更能找到顧客間共同的喜好，進而優化推薦系統(Huang, Chung, & Chen,

2004)；在社群偵測(Community Detection)方面，商品網路分析能夠找出商品隸屬於哪個社群，以此找出商品網路中的強關聯商品(Raeder & Chawla, 2009; Meng, Chaokun, Jeffrey & Jun, 2015)；在商品類別管理方面，商品網路分析賦予商品屬性變數如商品類別、購買此商品之顧客特質，以此可以幫助回答哪些商品同屬相同類別、哪種商品容易促發跨類別銷售等問題(Zinoviev, Zhu, & Li, 2015)。

然而雖然目前以交易商品為社群節點的商品網路分析有一定的研究，但不論是傳統的關聯規則還是新的商品網路分析依然對商品彼此併買缺乏解釋力，所以本研究希望利用以商品網路分析為基礎之指數隨機圖模型(Exponential Random Graph Modeling)來預測及解釋商品彼此併買的可能原因，幫助零售業者對交易資料有更好的分析應用。



### 第三章 模型介紹

指數隨機圖模型(Exponential Random Graph Modeling, ERGM)是一個用來檢測隨機圖或是網路圖模型中彼此連結關係模式的工具。相較於傳統對觀察網路(Observed Network)的敘述性分析如：密集度與中心性，ERGM 做為一種統計方法能夠更精確的描述觀察網路中的連結並非出於隨機，而是有許多結構特徵(Structural Feature)來影響網路結構(Snijders, 2011a; Rivera, Soderstorm, & Uzzi, 2010)。其中結構特徵包含：

1. 不均勻的分佈程度(Nonuniform Degree Distribution)：表示網路圖裡的節點不是每個都有一樣的連結傾向。舉例來說，以一個社團的朋友連結為一網路圖，在一個社團裡就有某些人比較容易結交到朋友。
2. 同質性(Homophily)：表示比起隨機機率，節點具有相同的特徵更容易彼此連結。舉例來說，以學校朋友連結圖為一網路圖，同樣參與相同社團的學生更容易成為朋友。
3. 傳遞性(Transitivity)：表示比起隨機機率，節點具有共同的連結節點彼此間更容易連結。舉例來說，以學校朋友連結圖為一網路圖，學生 A 與學生 B 是朋友，學生 C 與學生 A 是朋友，學生 B 與學生 C 是朋友的機率比較高。

ERGM 用以解釋或預測網路圖模型中節點連結的關係，其中網路圖模型中的簡單隨機圖型(Simple Random Graph)為原始的一般式，而本研究採用二元依賴模型(Dyadic Dependence Model)。簡單隨機圖型是從一個網路圖中隨機取出 n 個節點，使得其每個節點之間的連結機率一樣、節點彼此之間互相獨立的網路架構。相較於簡單隨機圖型，二元依賴模型將結構特徵考慮進去，使得節點彼此連結的機率不會一樣，節點彼此之間不互相獨立。

ERGM 可用以下原始一般式來解釋簡單隨機圖型的連結特徵(Erdős & Rényi, 1959; Frank & Strauss, 1986; Karonski, 1982)：

$$P(Y=y) = \left(\frac{1}{c}\right) \exp\{\theta L(y)\} \quad , \quad \text{其中}$$

- Y 是網路架構中隨機的關係（節點之間有連或沒連）
- y 是網路架構中觀察到的關係
- c 是一個常數用來確保觀察網路的機率落在 0 與 1 之間
- L(y)是網路 y 的邊(Edge)之數量
- $\theta$  是網路裡邊的係數值

本研究利用 ERGM 之二元依賴模型，可用以下公式來解釋連結特徵(Hunter, Handcock, Butts, Goodreau, & Morris, 2008):

$$\text{logit}(P(Y_{ij} = 1 | n \text{ actors}, Y_{ij}^c)) = \sum_{k=1}^k \theta_k \delta_{z_k(y)} \quad , \quad \text{其中}$$

- k 是模型中網路統計量的數量
- $\theta_k$ 是網路特徵的係數值
- $\delta_{z_k(y)}$ 是代表影響節點 i 與 j 之間是否連結的網路特徵
- $Y_{ij}^c$  代表所有二元節點除了  $Y_{ij}$
- $Y_{ij} = 1 | n \text{ actors}$  代表  $Y_{ij} = 1$  的機率是在其餘網路的條件下

ERGM 可以說是網路版的邏輯式迴歸(Logistic Regression)，為了更方便解讀，將上面的公式重新調整可用以下公式表示：

$$P(Y_{ij} = 1 | n \text{ actors}, Y_{ij}^c) =$$

$$\text{logistic}(\theta_1 \cdot \delta_{z_1(y)} + \theta_2 \cdot \delta_{z_2(y)} + \dots + \theta_n \cdot \delta_{z_n(y)})$$

$$\text{其中邏輯式函數為：} \frac{e^x}{1+e^x}$$

本研究將原始交易資料轉為基於商品網路分析下的網路架構，利用 ERGM 賦予每個商品屬性變數，透過模型結果預測及解釋商品網路圖中彼此連結的機率，以以下公式為例：

$$P(Y_{ij} = 1 | n \text{ actors}, Y_{ij}^c) = \frac{\exp\{\theta_0 + \theta_1 \cdot X_1 + \theta_2 \cdot X_2 + \dots + \theta_n \cdot X_n\}}{1 + \exp\{\theta_0 + \theta_1 \cdot X_1 + \theta_2 \cdot X_2 + \dots + \theta_n \cdot X_n\}}, \text{ 其中}$$

- $Y_{ij}=1$  代表商品 i 與 j 之間同時購買
- $\theta$  是網路特徵的係數值
- $X$  是商品屬性變數，例如：金額、類別



## 第四章 ERGM 模型分析結果

### 第一節 資料敘述

本研究以台灣知名超商 2017 年之商品交易資料實作 ERGM，此超商截至今年在台灣有三千多間分店並且不斷擴張，其宗旨就是為顧客提供更多的附加價值，以大數據、電子商務及新型態超商為三大領域當作未來發展方向。

本研究取得此超商 2017 年南區 12 月份之商品交易資料，其中包含 111 間店及 6,122 種商品，原始資料為每次消費個別商品的交易紀錄，共取得 4,084,342 筆資料，選取包含「會員卡號」、「性別」、「年齡」、「交易 id」、「資料日期」、「資料時間」、「餐別帶」、「縣市別」、「店鋪代號」、「主商圈」、「品號-品名稱」、「群號-群名稱」、「單品名稱」、「銷售數量」、「銷售單價」、「交易金額」等十六項交易資料欄位。經過刪除原始資料中「單價金額」為“0”、「單品名稱」為空及超商提供需刪除之 1,290 種商品，處理後最後取得 3,120,908 筆資料。

由圖 2 可以看出其中五筆商品交易資料，當交易資料中的「交易 id」相同時，就視為是同張發票，同張發票內的商品則視為彼此併買，例如圖 2 中編號 3、4 的「交易 id」相同，代表商品「厚切豬排蛋三明治」與「光泉高鈣牛乳」同時被併買。

會員卡號	性別	年齡	交易id	資料日期	資料時間	餐別帶	縣市別
1 NA	NA	NA	01667620171201202614000119716102	2017/12/1	20:26:14	晚餐時間帶	高雄市
2 NA	NA	NA	01648720171201180250000118954202	2017/12/1	18:02:50	晚餐時間帶	高雄市
3 NA	NA	NA	01648720171201180659000112503101	2017/12/1	18:06:59	晚餐時間帶	高雄市
4 NA	NA	NA	01648720171201180659000112503101	2017/12/1	18:06:59	晚餐時間帶	高雄市
5 NA	NA	NA	01648720171201181114000118955302	2017/12/1	18:11:14	晚餐時間帶	高雄市
店鋪代號	主商圈	品號-品名稱	群號-群名稱	單品名稱	銷售數量	銷售單價	交易金額
1	11290	住宅型 58-香煙	581-國產淡煙	新樂園 1 號香菸	1	80	80
2	7970	文教型 29-冷藏飲料	294-冷藏豆米漿	統一陽光黃金豆漿	2	20	36
3	7970	文教型 10-三明治	106-三明治 4 度 C	厚切豬排蛋三明治	1	42	39
4	7970	文教型 28-其他乳品	282-小包裝調味乳	光泉高鈣牛乳	1	36	20
5	7970	文教型 29-冷藏飲料	294-冷藏豆米漿	義美古早傳統低糖豆奶	1	23	21

圖 2 部份商品交易資料

由表 1 可以看出交易資料中每張發票金額及每張發票購買商品件數之敘述統計，從金額來看，平均一張發票金額為 88 元，超商可以根據此金額制定特定行銷活動，例如當消費滿超過 88 元即可獲得一點，以此提高民眾平均消費金額；從購買商品之數量來看，基本上消費者至超商購物只買一件商品的比率很高。

表 1 每張發票金額及購買商品件數之敘述統計

	最小值	25%	中位數	平均	75%	最大值
每張發票金額	0	29	57	88	104	45,000
每張發票購買商品件數	1	1	1	1.99	2	836

由圖 3 可以看出各主商圈交易資料中的交易金額，能夠明顯看出「住宅型」主商圈交易金額遠大於其他主商圈。

主商圈類型	工業型	文教型	住宅型	車站型	娛樂型	商業型	景點型	幹道型	辦公型	醫院型
交易總金額	\$ 4,237,250	\$ 6,336,945	\$ 131,996,517	\$ 1,958,225	\$ 2,573,465	\$ 11,382,483	\$ 1,068,611	\$ 9,940,887	\$ 3,080,223	\$ 1,444,400

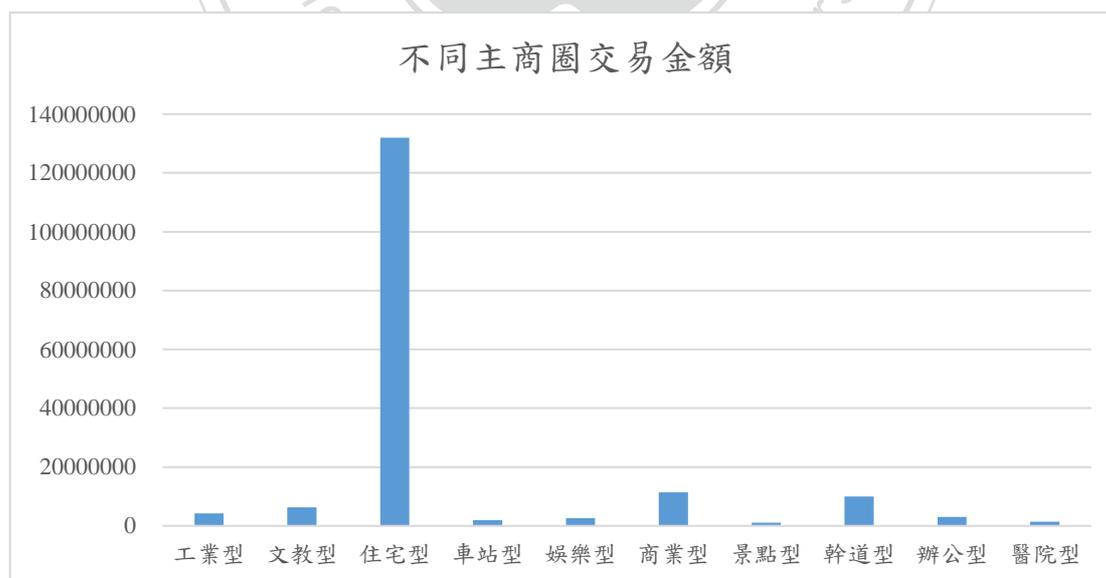


圖 3 各主商圈之交易金額

表 2 列出全部 3,120,908 筆資料中有哪幾種「品號-品名稱」及「群號-群名稱」，並列出「品號-品名稱」之交易筆數。

表 2 全部商品交易資料「品號-品名稱」與「群號-群名稱」統計

品號-品名稱	群號-群名稱	交易筆數
01-杯裝飲料	011-杯裝其它、012-杯裝飲料、 014-冰沙飲料、015-咖啡週邊商品	144,207
02-霜淇淋	021-霜淇淋	3,190
03-蒸箱食品	031-一般包子、032-蒸箱冷藏	7,142
04-燒烤食品	041-燒烤熱狗、042-蕃薯食品	48,356
05-關東煮	051-一般魚漿製品、053-微波關東煮、 054-蔬菜類製品、056-常溫商品	57,251
06-飯糰	061-飯糰、065-飯糰 4 度 C	87,417
07-壽司、手卷	071-壽司、072-手卷	17,563
08-便當	083-便當 4°C	17,810
09-微波主食	091-中式主食、092-西式主食、 093-日式主食、094-粥品、095-其他飯類	42,687
10-三明治	106-三明治 4 度 C、107-調理麵包 4 度 C	35,344
12-調理麵	121-涼麵 2 配、122-熱食麵 2 配、 123-熱湯麵 2 配、124-義大利麵	21,999
13-蔬菜水果	131-截切水果、132-冷藏原果（免稅）、 133-常溫原果（免稅）、135-葉菜蔬菜（免稅）、 136-根莖果（免稅）、137-截切蔬菜	11,265
14-蔬菜、沙拉	141-沙拉	9,306
15-小菜、微波	153-微波湯品、154-小菜、155-微波小吃	52,492
16-冷藏食品	161-豆腐、162-生麵 163-、火腿香腸、 164-奶油起司	9,191
17-冷藏甜點	171-布丁、172-優格、173-果凍、 175-常溫甜點、176-手工甜點、177-甜湯飲品	41,513
18-麵包	181-甜麵包、182-丹麥麵包、183-鹹麵包、 186-烘焙輕食	99,928
19-吐司蛋糕	191-小吐司餐包、192-蛋糕、193-家庭號吐司 餐包、195-小朋友麵包、196-常溫茶點	86,683

22-蛋品	221-洗選蛋、222-特殊蛋、224-加工蛋、 226-茶葉蛋	69,097
23-冰品	231-冰棒、232-杯冰、233-甜筒、234-家庭號冰 品、235-高級冰品、237-衛生冰塊	16,856
24-冷凍食品	241-冷凍調理、242-冷凍麵食、246-天和冷凍 食品－免稅、247-天和冷凍食品－應稅	22,191
27-鮮乳	271-大包裝全脂鮮乳、272-小包裝全脂鮮乳、 274-大包裝低脂鮮乳、275-小包裝低脂鮮乳	63,474
28-其他乳品	281-大包裝調味乳、282-小包裝調味乳、 283-冷藏植物奶、284-大包裝優酪乳、 285-小包裝優酪乳、286-發酵乳	151,604
29-冷藏飲料	291-冷藏茶飲料、292-冷藏果汁、293-冷藏咖 啡、294-冷藏豆米漿、295-其他冷藏飲料、 296-冷藏奶茶、297-冷藏常溫飲料	300,975
31-茶飲料	311-紅茶、312-綠茶、313-烏龍茶、314-奶茶、 315-其他茶、316-果茶	272,615
32-果汁飲料	321-100%果汁、322-未滿100%果汁、 323-蔬菜汁	17,696
33-嗜好飲料	331-咖啡飲料、332-碳酸飲料、333-其他嗜好 飲料、337-日翊配送飲料、 338-預購－嗜好飲料	98,038
34-健康飲料	341-機能飲料、342-運動飲料、343-水、 344-其他水	143,919
35-口香糖	351-一般口香糖、352-無糖口香糖、 353-兒童口香糖	28,125
36-巧克力	361-條狀巧克力、362-顆粒狀巧克力、363-盒 裝巧克力、364-兒童巧克力 365-其他巧克力	76,966
37-糖果	371-機能性糖果、372-一般糖果、 373-兒童糖果	53,698
38-點心零食	381-袋裝洋芋片、382-盒罐裝洋芋片、383-袋 裝零食、385-米果仙貝、386-其他零食	73,814
39-餅乾	391-夾心餅、392-威化餅、393-蘇打餅、 394-卷心餅、395-棒狀餅乾、396-煎薄派酥、 397-糕點類	26,640

40-珍味豆類	401-海產製品、402-肉類製品、403-堅果、 404-花生瓜子	13,670
41-蜜餞豆干	411-蜜餞、412-豆干	8,972
43-泡麵	431-大杯麵、432-小杯麵、433-大碗麵、 434-小碗麵 435-袋麵	77,337
44-即食食品	441-麵條、442-湯、443-調理包、444-粥、 447-日翊配送加工 1	3,735
45-瓶、罐頭	451-珍味瓶、452-魚罐頭、453-肉罐頭、 454-農產罐頭、455-果醬、奶油	4,168
46-調味品	461-食用油、462-醬油、463-調味醬、464-香 料、465-料理糖、味精、466-乾物、粉、 467-鹽	105
47-免稅食品	471-米、473-麵粉、 477-日翊配送加工（免稅）	19
48-營養食品	481-補品、482-健康食品、488-好市多商品	6,821
49-嗜好食品	491-咖啡、492-茶、493-奶精、煉乳、 494-即食糖、495-其他嗜好食品、496-觀光區 加工類食品、497-日翊配合加工 2	6,408
50-食品禮盒	501-常態果子禮盒、503-常態其他禮盒、 504-節慶果子禮盒、505-節慶補品禮盒	59
51-四大節慶 （母親節／ 中秋節）	511-原店常溫冷藏 1、 512-原店冷凍日翊常溫 1、516-宅配	38
52-四大節慶 （年菜／ 端午節）	521-原店常溫冷藏 1、 522-原店冷凍日翊常溫 1、523-他店連線、 525-預購訂購到店、526-宅配	242
53-預購－3C 商品／日用	531-日翊常溫 1、532-日翊常溫 2、537-宅配	32
55-預購－飲料	555-預購訂購到店	6
56-預購－零食	566-宅配、568-預購訂購到店（應稅）	98
57-文化出版品 預購	571-應稅出版品預購（店）、 572-免稅出版品預購（店）	20
58-香煙	581-國產淡煙、582-國產濃煙、584-進口淡 煙、585-進口濃煙、586-進口涼煙、587-雪茄	432,746

59-國產烈酒	591-國產料理酒、592-國產蒸餾酒、 593-國產發酵酒、594-國產水果酒	19,147
60-進口烈酒	601-進口威士忌、602-進口白蘭地、 603-進口酒禮盒、604-進口水果酒、 605-進口清酒、606-進口白酒、607-進口調酒	7,973
61-啤酒	611-國產啤酒、612-進口啤酒、613-國產瓶 啤、614-進口瓶啤、616-即飲調酒	48,188
63-其他酒類	631-國產藥酒	568
65-預購一日用 ／流行音樂	655-訂購到店商品	2
66-美妝用品	661-彩妝用品、662-護膚保養品、 663-美髮用品、664-美妝雜貨、665-流行飾 品、667-刮鬍用品	4,345
67-盥洗用品	671-洗髮用品、672-沐浴用品、673-洗臉用 品、674-口腔牙膏、675-牙刷	2,778
68-保健衛生	681-兩性用品、682-醫療用品、683-保健用 品、685-醫院區巡迴商品、686-醫院區用品	15,050
69-棉、紙製品	691-衛生紙、692-衛生棉、694-紙尿褲	19,138
70-紡織品	701-毛巾、702-男內衣褲、705-絲襪、 706-棉襪、707-日翊紡織品、 708-預購－紡織品	1,247
71-家庭用品	711-衣物洗滌劑、712-家庭清潔劑、 713-殺蟲防蟲劑	67
72-家庭雜貨	721-餐具用品、722-家雜用品、723-電器五 金、725-學區型商品、726-其他節令商品	102
73-寵物、園藝	731-狗食	7
75-文具	751-益智文具、752-一般文具、753-筆	16
76-玩具	761-一般玩具、762-日翊玩具、 763-巡迴玩具、764-檔期玩具	230
77-藥局複合測 試	771-藥局複合食品、776-美容生活用品	9
78-電池	783-特殊電池	3
79-個人用品	791-雨具用品、796-觀光區用品	24
80-卡類商品	802-電子票證（交通票卡）、803-電信預付卡	266

81-虛擬點卡	811-虛擬電信卡、813-虛擬遊戲卡、 814-GIFTCARD 虛擬卡、815-虛擬其他卡、 816-一元 PIN、817-一元 PIN 卡、 818-一元 PIN 錢	26,373
82-全家行動購 冷凍食(鹹)	821-免稅宅配	1
83-文創娛樂	831-金光布袋戲、832-霹靂布袋戲、 837-預購-店配、838-文創商品預購(店)	10,738
84-報紙	841-一般日報、842-一般晚報、843-財經報、 844-外文報	113,718
85-應稅出版品	851-電腦資訊、852-休閒育樂、授權週邊、 854-漫畫單行本、855-休閒育樂、授權週邊、 856-其他叢書、地圖旅遊、857-文創商品	640
86-免稅雜誌	861-週刊、862-男性時尚、863-女性時尚、 864-電腦資訊、865-流行情報、 866-語言汽車、867-漫畫期刊、868-其他雜誌	3,789
87-應稅單行本	872-休閒育樂、授權週邊、873-小說勵志、 874-漫畫單行本、876-其他叢書、地圖旅遊、 877-其他書籍、878-其他叢書	537
88-視聽娛樂	881-服務商品、883-3C 週邊、886-影音光碟	2,953
89-資訊商品	891-娛樂票券、892-資訊開卡包休閒類、 894-GIFTCARD 實體卡(銷售)、895-上網 卡、896-娛樂票券 2、897-資訊服務手續費	29,873
91-服務	911-全錄影印、914-宅配、915-國際快遞、 916-服務手續費、917-服務機台	48,394
92-服務	924-宅配	629
93-代售商品 (一)	931-郵票、933-商品預售應稅、 937-商品預售免稅	91
94-預購-酒類	941-日翊常溫、945-訂購到店商品	99
97-全店暨聯促 贈品	977-行銷部贈品、978-行銷部贈品(預換)	23
98-代售商品 (二)	983-代售票券、984-代售票券 2、986-特殊店代 售、987-國際電話卡代售、 988-娛樂免稅票券	362

## 第二節 資料分析結果

本研究利用 R 語言設計一套 ERGM 流程，以此回答國內知名超商商品併買的各项問題，使用工具為 R 語言中 library('statnet')及 library('ergm')。圖 4 的流程架構圖將 ERGM 流程共分為以下四大階段，詳細各階段流程於文字描述。

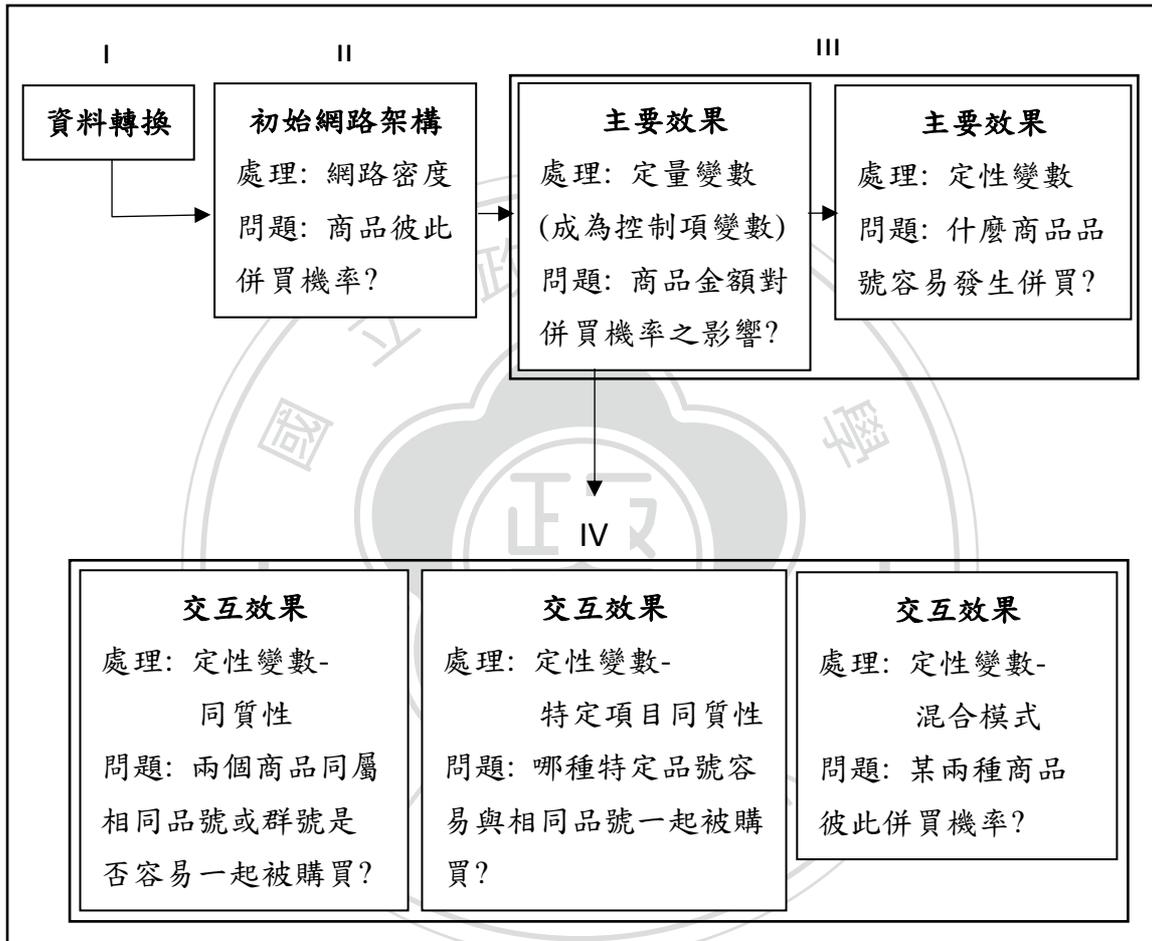


圖 4 流程架構圖

### 階段一：資料轉換

為了利用 ERGM 來檢測處理後交易資料之連結關係，須將處理後資料從原始的表格型數據結構(Data Frame)轉為網路架構(Network)之形式，以圖 5 為範例，編號 1、2、3、4 擁有相同「交易 id」，也就是這個交易 id 裡面的商品彼此視為

併買，當轉為網路架構時會在不考慮銷售數量的情況下，將這四項商品進行兩兩排列組合，在網路架構下商品彼此併買視為一條連線(edge)，所以可以得出六條連線，同時必須排除在同個「交易 id」下只買一項商品的交易資料，也就是編號 5 跟 6 的交易資料會被排除在外。由圖 6 可以看到經過上述處理得到初始網路架構之資訊，包括 4,564 個節點（也就是 4,564 項商品）、582,210 條線（也就是有 582,210 種併買發生）、密集度 0.0559 之不具方向性網路圖，同時為了預測及解釋商品網路圖中彼此連結的機率，創立三種節點屬性變數，包含「商品金額」、「商品所屬品號-品名稱」及「商品所屬群號-群名稱」，以下表 3 呈現「商品金額」之敘述統計，表 4 及表 5 分別呈現品號及群號出現次數前五高之敘述統計。

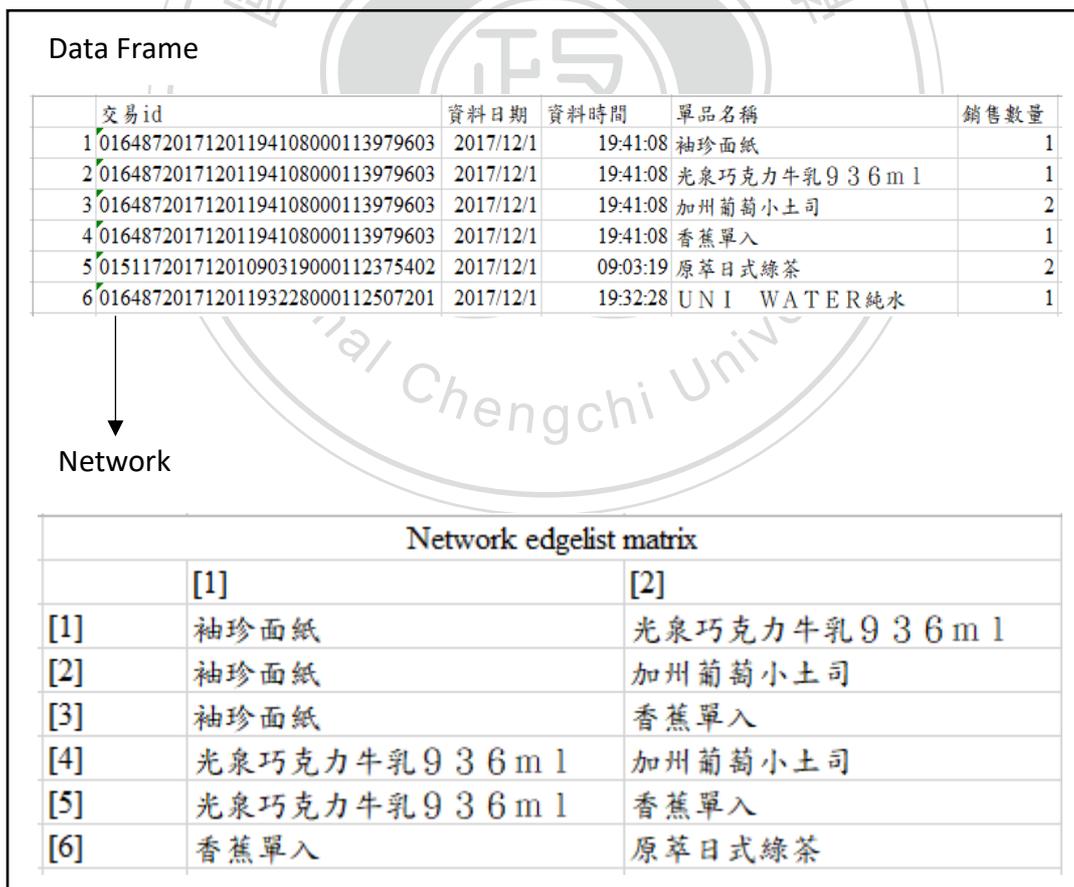


圖 5 資料轉換範例圖

```

Network attributes:
vertices = 4564
directed = FALSE
hyper = FALSE
loops = FALSE
multiple = FALSE
bipartite = FALSE
total edges = 582210
missing edges = 0
non-missing edges = 582210
density = 0.0559131

```

圖 6 網路架構資訊

表 3 節點屬性變數「商品金額」之敘述統計

	最小值	25%	中位數	平均	75%	最大值
商品金額	2	36	75	154	149	5988

表 4 節點屬性變數「商品所屬品號-品名稱」之敘述統計

品名稱	58- 香煙	86- 免稅雜誌	81- 虛擬點卡	36- 巧克力	37- 糖果
出現次數	299	191	186	164	163

表 5 節點屬性變數「商品所屬群號-群名稱」之敘述統計

群名稱	584- 進口淡煙	883- 3 C 週邊	813- 虛擬遊戲卡	862- 男性時尚	365- 其他巧克力
出現次數	139	138	106	85	82

## 階段二：初始網路架構

在一開始零售業者若想要了解在商品交易資料中，商品與商品彼此併買機率為何，在初始網路架構中只考慮兩個節點是否有連結，也就是有幾條「邊」來解釋網路架構之密度，圖 7 顯示初始網路架構資訊，可以利用估計值計算出勝算比取自然對數(Log Odds)之結果，其中 Log Odds 公式為：

$$\log(\text{Odds}) = \log\left(\frac{P_i}{1-P_i}\right) = \alpha + \beta * x$$

因為這裡只考慮「邊」的估計值，所以其 Log Odds 計算為：

$$\begin{aligned}\log(\text{Odds}) &= (-2.826420) * \text{edge} \\ &= (-2.826420) * 1\end{aligned}$$

得出任何一個邊出現的 Log Odds 為-2.826420，由於在網路架構中，每發生一次併買都對整體網路的連結量增加數量 1，所以 edge 為 1，而其對應的機率即為  $\exp(-2.826420)/(1 + \exp(-2.826420)) = 0.0559$ ，跟圖 4 所得到的密集度資訊相同，也就是超商中商品交易網路圖商品彼此併買機率約為 0.0559。

```

Monte Carlo MLE Results:
      Estimate Std. Error MCMC % p-value
edges -2.826420   0.001349     0 <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 14435159 on 10412766 degrees of freedom
Residual Deviance: 4489379 on 10412765 degrees of freedom

AIC: 4489381    BIC: 4489395    (Smaller is better.)

```

圖 7 初始網路架構模型結果

### 階段三：主要效果(Main Effects)

主要效果幫助零售業者了解不同商品屬性對商品彼此連結機率影響為何，本研究以商品金額及商品品號這兩種商品屬性為例，在得知併買機率之後，零售業者若欲了解商品金額越高，對商品彼此併買機率是提高還是下降，或是若欲了解哪種商品品號比較容易跟同品號或是其他品號之商品發生併買行為，這兩個問題在 ERGM 的架構下皆透過主要效果探討個別節點本身之商品屬性對 Log Odds 所造成之影響，分別用以下兩點回答。

1. 利用 ERGM 中 nodecov 函數處理主要效果裡節點屬性(Attribute)中定量變數 (Quantitative Variable)之資料，其中商品交易資料中「商品金額」之節點屬性變數就屬於定量變數。圖 8 顯示初始網路加入 nodecov 函數後之結果，可以

初步看到「商品金額」負的估計值-0.01524 意味著「商品金額」這個節點屬性變數對商品彼此併買機率有負面效果，而整體估計值所計算出兩個商品之間連結 Log Odds 為：

$$\begin{aligned} \log(\text{Odds}) &= (-0.6009) * \text{edge} + (-0.01524) * \text{Price} \\ &= (-0.6009) * \text{edge} + (-0.01524) * (\text{節點 i 的商品金額} \\ &\quad + \text{節點 j 的商品金額}) \\ &= (-0.6009) * 1 + (-0.01524) * \text{兩個節點商品金額之總額} \end{aligned}$$

舉例來說當兩個商品的連結用全部商品之最低金額來計算，其 Log Odds 為：

$$\begin{aligned} \log(\text{Odds}) &= (-0.6009) + (-0.01524) * (2+2) \\ &= -0.66186 \end{aligned}$$

若是兩個商品金額用全部商品之平均金額來計算，其 Log Odds 為：

$$\begin{aligned} \log(\text{Odds}) &= (-0.6009) + (-0.01524) * (154+154) \\ &= -5.29482 \end{aligned}$$

而其對應的機率分別約為 0.34032 及 0.00499，可以看出當商品金額越高，在商品網路架構裡兩個商品彼此連結機率越低。此外「商品金額」這個節點屬性變數會成為之後階段各模型的控制項變數，也就是將「商品金額」這個變數對併買的影響可能排除。

```

Monte Carlo MLE Results:
      Estimate Std. Error MCMC % p-value
edges          -6.009e-01  3.071e-03    0 <1e-04 ***
nodecov.dollar -1.524e-02  2.624e-05    0 <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 14435159 on 10412766 degrees of freedom
Residual Deviance: 3723093 on 10412764 degrees of freedom

AIC: 3723097    BIC: 3723125    (Smaller is better.)
    
```

圖 8 加入 nodecov 函數之網路架構模型結果

2. 另一個處理 ERGM 主要效果中節點屬性的函數為 nodefactor 函數，負責定性變數(Categorical Variable)，其中商品交易資料中「商品所屬品號-品名稱」及「商品所屬群號-群名稱」之節點屬性變數就是屬於定性變數。為了回答哪種商品品號比較容易跟同品號或是其他品號之商品發生併買行為，圖 9 顯示在考慮商品金額這個控制項變數後，再加入「商品所屬品號-品名稱」之 nodefactor 函數後的模型結果，其中各商品品名稱的估計量會根據基準(base)品號進行比較，而此網路架構中商品品名稱的基準為「01-杯裝飲料」這個品號，以下將以「02-霜淇淋」品號估計值為例計算出兩個商品之間連結 Log Odds 為：

$$\log(\text{Odds}) = (-1.311) * \text{edge} + (-0.006262) * \text{兩個節點商品金額總額} + (-1.539)$$

假設兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算：

$$\log(\text{Odds}) = (-1.311) * 1 + (-0.006262) * (75+154) + (-1.539)$$

得出 Log Odds 為-4.283998，換算其勝算(Odds)約為 0.01379，而基準品號「01-杯裝飲料」之 Log Odds 以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算：

$$\log(\text{Odds}) = (-1.311) * 1 + (-0.006262) * (75+154)$$

得出 Log Odds 為-2.744998，換算 Odds 約為 0.06425，可以得出「02-霜淇淋」

品號之於「01-杯裝飲料」品號之勝算比(Odds Ratio)約為  $0.215 \left( \frac{0.01379}{0.06425} \right)$ ，也

就是代表商品屬於「02-霜淇淋」品號和任何品號之商品發生併買機率相較於

「01-杯裝飲料」品號來的低。

Monte Carlo MLE Results:

	Estimate	Std. Error	MCMC %	p-value
edges	-1.311e+00	1.510e-02	0	< 1e-04 ***
nodecov.dollar	-6.262e-03	2.957e-05	0	< 1e-04 ***
nodefactor.big category.02-霜淇淋	-1.539e+00	2.967e-02	0	< 1e-04 ***
nodefactor.big category.03-蒸箱食品	-1.965e-01	1.937e-02	0	< 1e-04 ***
nodefactor.big category.04-燒烤食品	8.510e-01	1.322e-02	0	< 1e-04 ***
nodefactor.big category.05-關東煮	4.461e-01	1.133e-02	0	< 1e-04 ***
nodefactor.big category.06-飯糰	7.833e-01	1.128e-02	0	< 1e-04 ***
nodefactor.big category.07-壽司、手卷	8.028e-01	1.719e-02	0	< 1e-04 ***
nodefactor.big category.08-便當	1.083e+00	1.905e-02	0	< 1e-04 ***
nodefactor.big category.09-微波主食	1.006e+00	1.165e-02	0	< 1e-04 ***
nodefactor.big category.10-三明治	8.085e-01	1.278e-02	0	< 1e-04 ***
nodefactor.big category.12-調理麵	9.456e-01	1.360e-02	0	< 1e-04 ***
nodefactor.big category.13-蔬菜水果	-9.380e-01	1.335e-02	0	< 1e-04 ***
nodefactor.big category.14-蔬菜、沙拉	2.783e-01	1.587e-02	0	< 1e-04 ***
nodefactor.big category.15-小菜、微波	1.051e+00	1.076e-02	0	< 1e-04 ***
nodefactor.big category.16-冷藏食品	-1.487e-01	1.396e-02	0	< 1e-04 ***
nodefactor.big category.17-冷藏甜點	5.466e-01	1.038e-02	0	< 1e-04 ***
nodefactor.big category.18-麵包	9.591e-01	1.003e-02	0	< 1e-04 ***
nodefactor.big category.19-吐司蛋糕	9.048e-01	1.013e-02	0	< 1e-04 ***
nodefactor.big category.22-蛋品	9.121e-01	1.946e-02	0	< 1e-04 ***
nodefactor.big category.23-冰品	-1.036e+00	1.154e-02	0	< 1e-04 ***
nodefactor.big category.24-冷凍食品	-5.480e-01	1.108e-02	0	< 1e-04 ***
nodefactor.big category.27-鮮乳	9.128e-01	1.138e-02	0	< 1e-04 ***
nodefactor.big category.28-其他乳品	8.407e-01	9.119e-03	0	< 1e-04 ***
nodefactor.big category.29-冷藏飲料	6.445e-01	8.412e-03	0	< 1e-04 ***
nodefactor.big category.31-茶飲料	6.536e-01	8.574e-03	0	< 1e-04 ***
nodefactor.big category.32-果汁飲料	4.755e-01	1.396e-02	0	< 1e-04 ***
nodefactor.big category.33-嗜好飲料	2.626e-01	9.652e-03	0	< 1e-04 ***
nodefactor.big category.34-健康飲料	5.839e-01	9.535e-03	0	< 1e-04 ***
nodefactor.big category.35-口香糖	1.698e-02	1.133e-02	0	0.133841
nodefactor.big category.36-巧克力	-5.242e-02	8.723e-03	0	< 1e-04 ***
nodefactor.big category.37-糖果	-2.981e-01	8.888e-03	0	< 1e-04 ***
nodefactor.big category.38-點心零食	2.881e-01	8.570e-03	0	< 1e-04 ***
nodefactor.big category.39-餅乾	-2.341e-01	9.271e-03	0	< 1e-04 ***
nodefactor.big category.40-珍珠豆類	-5.226e-02	1.087e-02	0	< 1e-04 ***
nodefactor.big category.41-蜜餞豆干	-1.176e-01	1.305e-02	0	< 1e-04 ***
nodefactor.big category.43-泡麵	1.087e-01	8.868e-03	0	< 1e-04 ***
nodefactor.big category.44-即食食品	-1.674e+00	1.602e-02	0	< 1e-04 ***
nodefactor.big category.45-瓶、罐頭	-1.222e+00	1.584e-02	0	< 1e-04 ***
nodefactor.big category.46-調味品	-3.718e+00	7.961e-02	0	< 1e-04 ***
nodefactor.big category.47-免稅食品	-3.720e+00	1.966e-01	0	< 1e-04 ***
nodefactor.big category.48-營養食品	-1.690e+00	1.393e-02	0	< 1e-04 ***
nodefactor.big category.49-嗜好食品	-1.211e+00	1.421e-02	0	< 1e-04 ***
nodefactor.big category.50-食品禮盒	-2.989e+00	1.802e-01	0	< 1e-04 ***
nodefactor.big category.51-四大節慶 (母親節/中秋節)	-3.433e+00	3.538e-01	0	< 1e-04 ***
nodefactor.big category.52-四大節慶 (年菜/端午節)	-2.310e+00	5.611e-02	0	< 1e-04 ***
nodefactor.big category.53-預購-3C商品/日用	-2.448e+00	4.479e-01	0	< 1e-04 ***
nodefactor.big category.55-預購-飲料	-4.231e+00	2.779e-01	0	< 1e-04 ***
nodefactor.big category.56-預購-零食	-3.497e+00	9.005e-02	0	< 1e-04 ***
nodefactor.big category.57-文化出版品預購	-4.053e+00	1.745e-01	0	< 1e-04 ***
nodefactor.big category.58-香煙	-2.177e-01	8.532e-03	0	< 1e-04 ***
nodefactor.big category.59-國產烈酒	-1.967e-01	1.782e-02	0	< 1e-04 ***
nodefactor.big category.60-進口烈酒	-8.134e-01	1.425e-02	0	< 1e-04 ***
nodefactor.big category.61-啤酒	-4.606e-01	1.006e-02	0	< 1e-04 ***
nodefactor.big category.63-其他酒類	-2.385e+00	6.259e-02	0	< 1e-04 ***
nodefactor.big category.66-美妝用品	-1.909e+00	1.546e-02	0	< 1e-04 ***
nodefactor.big category.67-盥洗用品	-1.588e+00	1.739e-02	0	< 1e-04 ***
nodefactor.big category.68-保健衛生	-9.096e-01	1.344e-02	0	< 1e-04 ***
nodefactor.big category.69-棉、紙製品	-7.002e-01	1.218e-02	0	< 1e-04 ***
nodefactor.big category.70-紡織品	-2.180e+00	2.745e-02	0	< 1e-04 ***
nodefactor.big category.71-家庭用品	-3.085e+00	7.439e-02	0	< 1e-04 ***
nodefactor.big category.72-家庭雜貨	-4.022e+00	7.214e-02	0	< 1e-04 ***
nodefactor.big category.73-寵物、園藝	-4.170e+00	4.476e-01	0	< 1e-04 ***
nodefactor.big category.75-文具	-5.046e+00	2.044e-01	0	< 1e-04 ***
nodefactor.big category.76-玩具	-3.629e+00	5.691e-02	0	< 1e-04 ***
nodefactor.big category.77-藥局複合測試	-3.675e+00	2.896e-01	0	< 1e-04 ***
nodefactor.big category.78-電池	-2.447e+00	3.025e-01	0	< 1e-04 ***
nodefactor.big category.79-個人用品	-4.527e+00	3.783e-01	0	< 1e-04 ***
nodefactor.big category.80-卡類商品	-4.845e+00	1.082e-01	0	< 1e-04 ***
nodefactor.big category.81-虛擬點卡	-7.611e-01	1.353e-02	0	< 1e-04 ***
nodefactor.big category.83-文創娛樂	-1.779e+00	1.840e-02	0	< 1e-04 ***
nodefactor.big category.84-報紙	-6.221e-02	1.710e-02	0	0.000274 ***
nodefactor.big category.85-應稅出版品	-3.700e+00	3.912e-02	0	< 1e-04 ***

nodefactor.big	category.86-免稅雜誌	-3.279e+00	2.320e-02	0	< 1e-04	***
nodefactor.big	category.87-應稅單行本	-3.506e+00	3.576e-02	0	< 1e-04	***
nodefactor.big	category.88-視聽娛樂	-2.543e+00	2.680e-02	0	< 1e-04	***
nodefactor.big	category.89-資訊商品	-1.768e+00	1.741e-02	0	< 1e-04	***
nodefactor.big	category.91-服務	-1.063e+00	1.541e-02	0	< 1e-04	***
nodefactor.big	category.93-代售商品(一)	-3.600e+00	1.567e-01	0	< 1e-04	***
nodefactor.big	category.94-預購-酒類	-2.850e+00	1.287e-01	0	< 1e-04	***
nodefactor.big	category.97-全店暨聯促贈品	-4.629e+00	2.003e-01	0	< 1e-04	***
nodefactor.big	category.98-代售商品(二)	-3.467e+00	1.160e-01	0	< 1e-04	***
---						
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						
Null Deviance: 14435159 on 10412766 degrees of freedom						
Residual Deviance: 3163148 on 10412684 degrees of freedom						
AIC: 3163312 BIC: 3164473 (Smaller is better.)						

圖 9 加入 nodefactor 函數之網路架構模型結果

勝算比(Odds Ratio)即為將兩種不同事件發生的勝算(Odds)用比值的方式來衡量兩組勝算之差異 ( $\frac{A Odds}{B Odds}$ )，在這個階段如果利用 Odds Ratio 當成衡量指標，相較於計算各別品號之機率，更能直覺且快速知道欲比較品號與基準品號誰的勝算較大或是較小，所以 Odds Ratio 的衡量指標只是提供給零售業者不同的方式來比較欲比較之結果，當然也能夠利用 Log Odds 換算成機率，利用機率比較連結機率大小。也因此零售業者可以根據各商品所屬品號與基準品號之 Odds Ratio 快速找出哪些品號之商品比較容易跟同品號或是其他品號之商品發生併買行為，表 6 的左表格可以看出商品品號相較於「01-杯裝飲料」基準品號勝算比大於一，代表更容易與其他商品發生併買，以及右表格中相較於「01-杯裝飲料」基準品號勝算比小於一的品號，則代表不容易與其他商品彼此併買。

表 6 部分品號與基準商品勝算比比較表

Odds Ratio < 1	勝算比	Odds Ratio < 1	勝算比
08-便當	2.95	75-文具	0.01
15-小菜、微波	2.86	80-卡類商品	0.01
09-微波主食	2.74	97-全店暨聯促商品	0.01
18-麵包	2.61	79-個人用品	0.01
12-調理麵	2.57	55-預購-飲料	0.01
27-鮮乳	2.49	73-寵物、園藝	0.02
22-蛋品	2.49	57-文化出版品預購	0.02
19-吐司蛋捲	2.47	72-家庭雜貨	0.02
04-燒烤食品	2.34	47-免稅食品	0.02
28-其他乳品	2.32	46-調味品	0.02

#### 階段四：交互效果(Interaction Effects)

交互效果幫助零售業者了解兩個商品屬性共同對商品彼此連結機率影響為何，也因此零售業者可以透過這個階段得知兩個商品同屬相同品號或群號是否容易一起被購買，甚至可以進一步知道哪種品號之商品最易在相同品號內發生併買行為，又或是零售業者想要了解特定兩個不同商品之併買機率。這三個問題主要透過在 ERGM 的架構下交互效果中的 nodematch 及 nodemix 函數來做分析，利用以下三點詳細說明三個問題之回答。

1. nodematch 函數為處理 ERGM 交互效果中節點屬性的函數之一，其主要目的是用來解釋商品彼此連結之機率是否與節點具同質性(Homophily)有關，商品具同質性相較於隨機機率，節點擁有相同特徵更容易彼此連結，因此我們可以透過「商品所屬品號-品名稱」及「商品所屬群號-群名稱」這兩種節點屬性變數來探討商品同屬相同品號或是群號是否較容易一起被購買。從圖 10 可以看到階段三加入 nodematch 函數後的結果，其中品號以「big category」表示；群號以「small category」表示，可以觀察到不論是品號還是群號其估計值皆

為正，代表兩個商品同屬相同品號或群號較容易一起被購買，對商品彼此併買機率有正面效果。以「商品所屬品號-品名稱」之估計值計算兩個商品之間連結 Log Odds 為：

$$\log(\text{Odds}) = (-0.6574) * \text{edge} + (-0.01527) * \text{兩個節點商品金額總額} + 1.257$$

若以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算：

$$\log(\text{Odds}) = (-0.6574) * 1 + (-0.01527) * (75+154) + 1.257$$

得到 Log Odds 為-2.89723，其對應機率為 0.05229，但在不考慮品號這個同質性的情況下，同樣以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算，其 Log Odds 為：

$$\log(\text{Odds}) = (-0.6574) * 1 + (-0.01527) * (75+154)$$

得到 Log Odds 為-4.15423，其對應機率為 0.01546，也就代表此網路架構中兩個商品具有相同品號之同質性時，相較於兩個商品分屬不同品號彼此併買之機率比較高。

```

Monte Carlo MLE Results:
      Estimate Std. Error MCMC % p-value
edges          -6.574e-01  3.091e-03    0 <1e-04 ***
nodecov.dollar  -1.527e-02  2.626e-05    0 <1e-04 ***
nodematch.big category  1.257e+00  7.472e-03    0 <1e-04 ***
nodematch.small category  5.494e-01  1.239e-02    0 <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 14435159 on 10412766 degrees of freedom
Residual Deviance: 3674930 on 10412762 degrees of freedom

AIC: 3674938    BIC: 3674994    (Smaller is better.)

```

圖 10 加入 nodematch 函數之網路架構模型結果

2. 雖然利用 nodematch 函數後可以知道商品同屬相同品號或是群號是否較容易一起被購買，然而卻不能知道各別品號或群號的詳細資訊，也就是到底哪種品號或群號之商品最易在相同品號或群號內發生併買行為，所以在 nodematch 函數中設定 diff = TRUE 時，就可以得到各別品號或群號之估計值。圖 11 顯示以「商品所屬品號-品名稱」利用 nodematch 函數加上 diff = TRUE 的結果，以「01-杯裝飲料」之估計值計算兩個商品之間連結 Log Odds 為例：

$$\log(\text{Odds}) = (-0.6612) * \text{edge} + (-0.01523) * \text{兩個節點商品金額總額} + 1.437$$

若以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算：

$$\log(\text{Odds}) = (-0.6612) * 1 + (-0.01523) * (75+154) + 1.437$$

得出 Log Odds 為-2.71187，其對應機率約為 0.06228，但在不考慮「01-杯裝飲料」品號這個同質性的情況下，以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算，其 Log Odds 為：

$$\log(\text{Odds}) = (-0.6612) * 1 + (-0.01523) * (75+154)$$

得到 Log Odds 為-4.14887，其對應機率為 0.0155，也就代表此網路架構中兩個商品同屬「01-杯裝飲料」品號之同質性時，彼此購買機率相較於兩個商品非同屬「01-杯裝飲料」品號來的高。其中某些品號之估計值為-Inf 代表此特定品號類別裡的商品數可能極少導致在此品號間沒有形成任何連結。

Monte Carlo MLE Results:

	Estimate	Std. Error	MCMC %	p-value
edges	-6.612e-01	3.151e-03	0	< 1e-04 ***
nodecov.dollar	-1.523e-02	2.689e-05	0	< 1e-04 ***
nodematch.big category.01-杯裝飲料	1.437e+00	5.492e-02	0	< 1e-04 ***
nodematch.big category.02-霜淇淋	1.250e+00	2.521e-01	0	< 1e-04 ***
nodematch.big category.03-蒸箱食品	2.471e+00	3.476e-01	0	< 1e-04 ***
nodematch.big category.04-燒烤食品	3.139e+00	2.210e-01	0	< 1e-04 ***
nodematch.big category.05-關東煮	2.828e+00	1.275e-01	0	< 1e-04 ***
nodematch.big category.06-飯糰	2.770e+00	1.234e-01	0	< 1e-04 ***
nodematch.big category.07-壽司、手卷	2.805e+00	4.374e-01	0	< 1e-04 ***
nodematch.big category.08-便當	2.999e+00	5.289e-01	0	< 1e-04 ***
nodematch.big category.09-微波主食	4.176e+00	1.544e-01	0	< 1e-04 ***
nodematch.big category.10-三明治	5.051e+00	4.168e-01	0	< 1e-04 ***
nodematch.big category.12-調理麵	4.416e+00	2.209e-01	0	< 1e-04 ***
nodematch.big category.13-蔬菜水果	5.220e-01	7.319e-02	0	< 1e-04 ***
nodematch.big category.14-蔬菜、沙拉	3.674e+00	3.164e-01	0	< 1e-04 ***
nodematch.big category.15-小菜、微波	5.314e+00	2.416e-01	0	< 1e-04 ***
nodematch.big category.16-冷藏食品	1.893e+00	1.227e-01	0	< 1e-04 ***
nodematch.big category.17-冷藏甜點	2.376e+00	7.553e-02	0	< 1e-04 ***
nodematch.big category.18-麵包	3.788e+00	1.245e-01	0	< 1e-04 ***
nodematch.big category.19-吐司蛋糕	3.060e+00	9.659e-02	0	< 1e-04 ***
nodematch.big category.22-蛋品	2.778e+00	5.382e-01	0	< 1e-04 ***
nodematch.big category.23-冰品	1.509e+00	3.542e-02	0	< 1e-04 ***
nodematch.big category.24-冷凍食品	1.638e+00	4.601e-02	0	< 1e-04 ***
nodematch.big category.27-鮮乳	3.225e+00	1.159e-01	0	< 1e-04 ***
nodematch.big category.28-其他乳品	2.210e+00	4.372e-02	0	< 1e-04 ***
nodematch.big category.29-冷藏飲料	1.577e+00	2.101e-02	0	< 1e-04 ***
nodematch.big category.31-茶飲料	1.532e+00	2.503e-02	0	< 1e-04 ***
nodematch.big category.32-果汁飲料	1.415e+00	1.873e-01	0	< 1e-04 ***
nodematch.big category.33-嗜好飲料	9.618e-01	4.817e-02	0	< 1e-04 ***
nodematch.big category.34-健康飲料	1.658e+00	4.809e-02	0	< 1e-04 ***
nodematch.big category.35-口香糖	1.478e+00	7.346e-02	0	< 1e-04 ***
nodematch.big category.36-巧克力	1.323e+00	2.070e-02	0	< 1e-04 ***
nodematch.big category.37-糖果	9.508e-01	2.097e-02	0	< 1e-04 ***
nodematch.big category.38-點心零食	1.591e+00	2.101e-02	0	< 1e-04 ***
nodematch.big category.39-餅乾	1.208e+00	2.763e-02	0	< 1e-04 ***
nodematch.big category.40-珍味豆類	2.215e+00	5.514e-02	0	< 1e-04 ***
nodematch.big category.41-蜜餞豆干	1.511e+00	1.051e-01	0	< 1e-04 ***
nodematch.big category.43-泡麵	1.626e+00	2.455e-02	0	< 1e-04 ***
nodematch.big category.44-即食食品	-7.309e-01	1.085e-01	0	< 1e-04 ***
nodematch.big category.45-瓶、罐頭	1.202e+00	8.175e-02	0	< 1e-04 ***
nodematch.big category.46-調味品	-1.278e+00	7.172e-01	0	0.074724
nodematch.big category.47-免稅食品	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.48-營養食品	4.730e-01	5.163e-02	0	< 1e-04 ***
nodematch.big category.49-嗜好食品	7.504e-01	7.417e-02	0	< 1e-04 ***
nodematch.big category.50-食品禮盒	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.51-四大節慶（母親節／中秋節）	1.084e+01	1.105e+00	0	< 1e-04 ***
nodematch.big category.52-四大節慶（年菜／端午節）	9.216e+00	1.096e-01	0	< 1e-04 ***
nodematch.big category.53-預購-3C商品／日用	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.55-預購-飲料	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.56-預購-零食	-3.934e-01	7.326e-01	0	0.591283
nodematch.big category.57-文化出版品預購	6.302e+00	4.708e-01	0	< 1e-04 ***
nodematch.big category.58-香煙	1.383e+00	1.672e-02	0	< 1e-04 ***
nodematch.big category.59-國產烈酒	4.415e+00	1.762e-01	0	< 1e-04 ***
nodematch.big category.60-進口烈酒	2.806e+00	7.600e-02	0	< 1e-04 ***
nodematch.big category.61-啤酒	1.852e+00	3.148e-02	0	< 1e-04 ***
nodematch.big category.63-其他酒類	2.107e+00	4.684e-01	0	< 1e-04 ***
nodematch.big category.66-美妝用品	1.282e+00	5.206e-02	0	< 1e-04 ***
nodematch.big category.67-盥洗用品	2.126e+00	6.999e-02	0	< 1e-04 ***
nodematch.big category.68-保健衛生	1.457e+00	7.289e-02	0	< 1e-04 ***
nodematch.big category.69-棉、紙製品	1.519e+00	5.705e-02	0	< 1e-04 ***
nodematch.big category.70-紡織品	1.766e+00	1.084e-01	0	< 1e-04 ***
nodematch.big category.71-家庭用品	-2.267e-01	1.010e+00	0	0.822488
nodematch.big category.72-家庭雜貨	-1.858e+00	5.052e-01	0	0.000234 ***
nodematch.big category.73-寵物、園藝	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.75-文具	-1.092e+00	1.021e+00	0	0.284985
nodematch.big category.76-玩具	1.020e-01	2.960e-01	0	0.730323
nodematch.big category.77-藥局複合測試	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.78-電池	-Inf	0.000e+00	0	< 1e-04 ***
nodematch.big category.79-個人用品	5.803e+00	1.201e+00	0	< 1e-04 ***
nodematch.big category.80-卡類商品	6.795e-01	2.341e-01	0	0.003704 **
nodematch.big category.81-虛擬點卡	3.407e+00	5.851e-02	0	< 1e-04 ***
nodematch.big category.83-文創娛樂	2.272e+00	6.001e-02	0	< 1e-04 ***
nodematch.big category.84-報紙	4.563e+00	7.182e-01	0	< 1e-04 ***
nodematch.big category.85-應稅出版品	-7.050e-01	1.600e-01	0	< 1e-04 ***

```

nodematch.big category.86-免稅雜誌 -3.310e-01 7.410e-02 0 < 1e-04 ***
nodematch.big category.87-應稅單行本 6.280e-01 1.223e-01 0 < 1e-04 ***
nodematch.big category.88-視聽娛樂 1.862e+00 1.084e-01 0 < 1e-04 ***
nodematch.big category.89-資訊商品 -9.281e-02 9.464e-02 0 0.326786
nodematch.big category.91-服務 2.008e-01 1.105e-01 0 0.069208 .
nodematch.big category.93-代售商品(一) 3.901e+00 7.424e-01 0 < 1e-04 ***
nodematch.big category.94-預購-酒類 1.005e+01 4.157e-01 0 < 1e-04 ***
nodematch.big category.97-全店豐聯促贈品 1.802e+00 6.530e-01 0 0.005772 **
nodematch.big category.98-代售商品(二) -Inf 0.000e+00 0 < 1e-04 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 14435159 on 10412766 degrees of freedom
Residual Deviance: 3664658 on 10412683 degrees of freedom

AIC: 3664824 BIC: 3665999 (Smaller is better.)
Warning: The following terms have infinite coefficient estimates:
nodematch.big category.47-免稅食品 nodematch.big category.50-食品禮盒 nodematch.big category.53-預購-3C商品/
日用 nodematch.big category.55-預購-飲料 nodematch.big category.73-寵物、園藝 nodematch.big category.77-藥局複合測
試 nodematch.big category.78-電池 nodematch.big category.98-代售商品(二)

```

圖 11 加入 nodematch 及 diff = TRUE 函數之網路架構模型結果

超商業者可以根據這個階段的結果找出哪些品號之商品容易在自己組內一起被購買，表 7 列出前 10 項容易與自己組內併買之品號。

表 7 前 10 項容易一起被購買之商品品號

排名	前 10 項容易一起被購買之品號	勝算比
1	51-四大節慶(母親節/中秋節)	51021.38
2	94-預購-酒類	23155.79
3	52-四大節慶(年菜/端午節)	10056.76
4	57-文化出版品預購	545.66
5	79-個人用品	331.29
6	15-小菜、微波	203.16
7	10-三明治	156.18
8	84-報紙	95.87
9	12-調理麵	82.76
10	59-國產烈酒	82.68

- 利用 nodematch 函數可以依據品號或是群號之同質性，得知哪些品號會容易在組內被一起購買，但卻無法了解特定商品 A 及特定商品 B 是否容易一起被購買，因為 nodematch 函數只能解釋節點是否具有同質性，也就是只能解釋「是」跟「否」兩個值，然而零售業者想要了解特定兩個不同商品之併買機

率，而 nodemix 函數適用於探討節點變數具有混合模式(Mixing Patterns)的值，例如：各個節點變數分屬 Yes、No 或 Others 之一的值，透過 nodemix 函數可以得到這三種不同值成對的估計值。nodemix 為處理 ERGM 交互效果中節點屬性的函數之一，主要負責定性變數，也因此零售業者可以根據 nodemix 函數得知兩種選定商品彼此連結之機率，舉例來說如圖 12 為超商業者於 12 月所提出的促銷活動，根據階段三的模式加入 nodemix 函數之後，將商品交易資料中屬於「A 商品名稱」的節點賦予其屬性「a」值；屬於「B 商品名稱」的節點賦予其屬性「b」值，將其他非屬於「A 商品名稱」及「B 商品名稱」的節點賦予其屬型「其他」值，透過 nodemix 函數會產生六種成對的估計值，分別為「a-a」、「a-b」、「b-b」、「a-其他」、「b-其他」、「其他-其他」，其中節點成對的估計值會根據基準進行 Odds Ratio 比較，可以自行設定基準，在這個範例以「其他-其他」為基準。從圖 13 可以看到在在基於「其他-其他」的組合下其他各成對的估計值，以「A 商品名稱」及「B 商品名稱」之估計值計算兩個商品之間連結 Log Odds 為：

$$\log(\text{Odds}) = (-0.6156) * \text{edge} + (-0.01520) * \text{兩個節點商品金額總額} + 1.686$$

若以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算：

$$\log(\text{Odds}) = (-0.6156) * 1 + (-0.01520) * (75+154) + 1.686$$

得出 Log Odds 為-2.4104，其對應機率為 0.08238，但相較於基準情況兩個商品節點屬性值屬於「其他」及「其他」，同樣以兩個商品金額一個用全部商品之中位數一個用全部商品之平均金額來計算，其 Log Odds 為：

$$\log(\text{Odds}) = (-0.6156) * 1 + (-0.01520) * (75+154)$$

得到 Log Odds 為-4.0964，其對應機率為 0.01636，可以明顯看出當兩個商品節點屬性值分屬於「a」及「b」時，兩個商品彼此連結機率比起「其他-其他」來的高，同時也比「a-b」、「a-其他」以及「b-其他」來的高，零售業者可以根據這階段的結果驗證當推出某種促銷活動時，某兩種商品組合的併買率高或低，是否有成功達到促銷活動的目的，或是若對某兩種品號之組合併買率有興趣，也能透過 nodemix 函數掌握連結機率。

編號	A商品名稱	面銷訊息	B商品名稱
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	究極巧克力泡芙
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	究極牛奶泡芙
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	雙色迷你泡芙（巧克力&牛奶）
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	辻利抹茶歐蕾布丁
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	義式提拉米蘇
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	靜岡抹茶塔
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	脆皮香草雙餡泡芙
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	伯爵巧克力餅乾泡芙
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	原味捲蛋糕
1203	杯裝飲料	活動貼標商品+咖啡，享貼標商品半價	巧克力捲蛋糕

圖 12 促銷活動內容

```

Monte Carlo MLE Results:
              Estimate Std. Error MCMC % p-value
edges                -6.156e-01  3.093e-03    0 <1e-04 ***
nodecov.dollar      -1.520e-02  2.626e-05    0 <1e-04 ***
mix.編號1203.a.a     1.524e+00  6.256e-02    0 <1e-04 ***
mix.編號1203.a.b     1.686e+00  9.215e-02    0 <1e-04 ***
mix.編號1203.b.b     4.790e+00  7.235e-01    0 <1e-04 ***
mix.編號1203.a.其他  1.364e-01  7.966e-03    0 <1e-04 ***
mix.編號1203.b.其他  6.537e-01  1.365e-02    0 <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 14435159 on 10412766 degrees of freedom
Residual Deviance: 3720006 on 10412759 degrees of freedom

AIC: 3720020    BIC: 3720119    (Smaller is better.)

```

圖 13 加入 nodemix 函數之網路架構模型結果

## 第五章 結論

對於零售業者來說，有關消費者之購物傾向的數據是他們珍貴的資料，透過傳統關聯規則的購物籃分析，可以提供零售業者更好的決策選擇與制定，然而傳統的購物籃分析無法解釋商品彼此併買的潛在原因為何，縮限了零售業者對商品交易資料的分析與應用，因此本研究以ERGM二元依賴模型(Hunter, Handcock, Butts, Goodreau, & Morris, 2008)來解釋商品網路圖中的連結（商品彼此併買）是受到什麼結構特徵（例如：商品金額、商品所屬品號）所影響。

本研究透過設計一套ERGM二元依賴模型流程，實作於國內知名連鎖便利商店之商品交易資料，對商品彼此併買進行分析，其流程共分為四個階段：第一階段對資料進行資料轉換，將商品交易資料從原始的Data Frame結構轉為Network架構，以此能利用ERGM得到初始商品網路圖；在第二階段利用ERGM得到網路架構之密集度，也就是超商交易資料中商品彼此之間併買機率；第三階段主要幫助零售業者了解商品屬性對商品連結機率之影響，探討ERGM主要效果中定量變數及定性變數，本研究例子結果顯示商品金額（定量變數）對商品彼此併買率有負面效果，也就是說當某項商品其金額越高，越不容易與其他商品發生併買行為，同時商品品號及群號（定性變數）也是屬於影響商品彼此併買的原因之一，零售業者可以得出哪些品號之商品易於跟同品號或是其他品號之商品發生併買行為，幫助零售業者對行銷之決策制定有更好的幫助；最後的第四階段討論兩種商品屬性共同影響商品連結機率，也就是ERGM交互效果對商品網路圖的影響，本研究例子結果顯示整體而言當網路架構中兩個商品具有相同品號或群號之同質性時，相較於兩個商品分屬不同品號或群號彼此併買之機率較高，零售業者也能同時得知哪種品號或群號之商品容易在自己組內一起被購買，也就是能夠瞭解哪些品號

或群號在組內是屬於強連結，此外零售業者也能掌握特定兩項商品或是哪兩種品號彼此連結之機率，幫助零售業者驗證他們感興趣之商品併買機率。

值得一提的是，本研究流程將「商品金額」設定為控制項變數，也就是在控制「商品金額」這個變數後，各個模型對商品彼此併買之機率的影響，但若不控制「商品金額」這個變數時，以流程階段三中商品品號及群號（定性變數）對商品網路之影響為例，一樣以「02-霜淇淋」品號及基準品號「01-杯裝飲料」分別算其機率，這時候將公式裡的金額考慮各自的商品金額，以「02-霜淇淋」品號之平均金額為27元，以及「01-杯裝飲料」品號平均金額為54元後，得出各自勝算分別為0.0305及0.1202，勝算比約為 $0.254 \left( \frac{0.0305}{0.1202} \right)$ ，與原始得出的勝算比結果不同，但結果解釋相同，也就是商品屬於「02-霜淇淋」品號和任何品號之商品發生併買機率相較於「01-杯裝飲料」品號來的低。本研究以控制「商品金額」為例計算各流程之Log Odds，目的是想要排除商品金額對併買之影響，只考慮各階段不同變數對併買機率所帶來的改變，但若零售業者想要將「商品金額」這個變數考慮進去，也能得到不同結果之研究。

本研究ERGM流程四個階段提供零售業者更多商品交易資料的分析與應用，而這個模型也能做到其他分析，因此未來的研究方向包含像是針對不同時間帶（早、午、晚餐帶）或是不同商圈型態（如：住宅、景點、商業）等等資料做分析，若是能將資料根據不同維度做分群分析，或許能夠提供更多有用的資訊給零售業者，同時未來研究之延伸方向包含能以貨架別為單位做分析，也就是能夠了解目前超商商品陳設位置是否安排妥當，透過對貨架別分析期望改變商品銷售量。

透過這個流程同時能幫助零售業者了解商品連結的可能原因，不只可以對零售交易資料之分析有更進一步認識，也能幫助行銷人員應用在行銷、促銷活動上，例如更主動地將可能發生併買之商品推出併買活動，以此提高營業額。

## 參考文獻

- Agrawal, R., Imielinski, T., & Swami, A. (1993). Database mining: A performance perspective. *IEEE Transactions on Knowledge and Data Engineering*, 5(6), 914–925.
- Agrawal, R., & Srikant, R. (1994). Fast algorithm for mining association rules in large database. Research Report RJ 9839, IBM Almaden Research Center, Santiago, Chile.
- Akter, S. & Fosso, Wamba, S. (2016). Big Data Analytics in E- Commerce: A Systematic Review and Agenda for Future Research.
- Balkundi, P., & Kilduff, M. (2006). The ties that lead: A social network approach to leadership. *The Leadership Quarterly*, 17, 419-439.
- Bonchi, F., Castillo, C., Gionis, A., & Jaimes, A. (2011). Social network analysis and mining for business applications. *ACM Transactions on Intelligent Systems and Technology*, 2, 3, Article 22.
- Chen, Y. L., Tang, K., Shen, R. J., & Hu, Y. H. (2005). Market basket analysis in a multiple store environment. *Decision Support Systems*, 40(2), 339–354.
- Chiu, C., Ku, Y., Lie, L., & Chen, Y. (2011). Internet auction fraud detection using social network analysis and classification tree approaches. *International Journal of Electronic Commerce*, 15(3), 123–147.
- Coleman, J., Menzel, H., & Katz, E. (1966). *Medical Innovations: A Diffusion Study*. Bobbs Merrill.
- Erdős, P., & Rényi, A. (1959). On Random Graphs, I. *Publicationes Mathematicae(Debrecen)*, 6, 290-297.
- Frank, O., Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81(395), 832–842.
- Freeman, L.C. (1997). A set of measures of centrality based on betweenness. *Sociometry*, 40, 35–41.
- Huang, Z., H. Chen, D. Zeng. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Intelligent Systems and Technology*, 22(1), 116–142.

- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). Ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3), 1-29.
- Jin, K. (2013). Social Network Analysis of Facebook Brand Communities. Saint Mary's University, Halifax, Nova Scotia. Research Project for Degree of Business Administration, Saint Mary's University.
- Karonski, M. (1982). A review of random graphs. *Journal of Graph Theory*, 6(4), 349-389.
- Kaur, M., Kang, S. (2016). Market Basket Analysis: Identifying the changing trends of market data using association rule mining, International conference on Computational Modeling and Security. *Procedia Computer Science*, 78-85.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Alsteyne, M. V. (2009). Computational social science. *Sci*, 323, 5915, 721–723.
- Meng, W., Chaokun, W., Jeffrey, X. Y., Jun, Z. (2015). Community detection in social networks: an in-depth benchmarking study with a procedure-oriented framework. *Proceedings of the VLDB Endowment*, v.8 n.10, p.998-1009.
- Mitchell, J. C. (1969). The Concept and Use of Social Networks. Pp. 1-50 in *Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns*, edited by J. Clyde Mitchell. Manchester, England: Manchester University Press.
- Mostafa, M. (2015). Knowledge discovery of hidden consumer purchase behavior: a market basket analysis IJDATS, 7 (4) (2015), pp. 384-405.
- Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *J. Information Science*, 28, 441-453.
- Qi, X., Fuller, E., Wu, Q., Wu, Y., & Zhang, C.-Q. (2012). Laplacian centrality: A new centrality measure for weighted networks. *Information Science*, 194, 240–253.
- Raeder, T., Chawla, N. V. (2009). Modeling a store's product space as a social network. In *Proceedings of the 2009 international conference on advances in social network analysis and mining*. Athens, Greece, pp. 164–169.
- Rivera, M. T., Soderstrom, S. B., & Uzzi, B. (2010). Dynamics of dyads in social network: Assortative, relational, and proximity mechanisms. *Annual Review of Sociology*, 36, 91-115.

- Scott, John. (1991). *Social network analysis: A handbook*. London: Sage.
- Snijders, T. A. B. (2011a). Statistical models for social networks. *Annual Review of Sociology*, 37, 131-153.
- Tewari, A.S., Kumar, A., & Barman, A.G. (2014). Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining. *Advance Computing Conference (IACC), IEEE International*, 500 - 503.
- Wasserman, S., & Faust, K. (1994). *Social network analysis*. Cambridge, MA: Cambridge University Press.
- Wasserman, S., Pattison, P.E. (1996). Logit models and logistic regression for social networks. I. An introduction to Markov graphs and  $p^*$ . *Psychometrika*, 61(3), 401–425.
- Watts, D. J. (2004). The “new” science of networks. *Ann. Rev. Sociol.* 30, 243–270.
- Wellman, B., & Berkowitz, S. D (Eds.). (1988). *Social structures: A network approach*. Cambridge: Cambridge University Press.
- Zinoviev D., Zhu Z., Li K. (2015). Building mini-categories in product networks. *In Complex Networks VI*. Vol. 597. Springer, Cham.