

## Sulle orme del consumatore: *affinity analysis* e *knowledge visualization* per il processo decisionale nella distribuzione commerciale

Letizia Lo Presti<sup>1</sup>, Vittoria Marino<sup>2</sup>, Paolo Di Betta<sup>3</sup>

### Abstract (max 250 parole)

L'articolo analizza i reali comportamenti di acquisto dell'acquirente all'interno del punto vendita e individua i *pattern* latenti più significativi attraverso un'analisi di affinità di 128.364 scontrini estratti da un dataset di 715.662 transazioni effettuate in un anno. Lo studio individua le strutture associative più ricorrenti relativamente alla singola merceologia. Attraverso la tecnica della *Market Basket Analysis* è stato possibile individuare le categorie merceologiche che creano "poli d'attrazione" sul punto vendita stimolando acquisti di altre categorie. In particolar modo, applicando le regole associative sul punto vendita è stato possibile identificare le "aree calde" e le "aree fredde". L'articolo evidenzia quindi alcune riflessioni sulle strategie che possono adottare gli store manager.

**Parole chiave:** Market Basket Analysis; layout; comportamento d'acquisto; knowledge visualization; affinity analysis

### Obiettivo della ricerca

Il cambiamento nella struttura dei consumi delle famiglie, orientata sempre più verso una programmazione degli acquisti e verso l'e-commerce, ha cambiato il modo di fare shopping e, per certi versi, razionalizzato la spesa (sono molti i consumatori che si recano presso il punto vendita dopo aver fatto delle ricerche online) minacciando l'idea stessa di *shopping experience*, intesa in senso tradizionale. La quantità di informazioni a disposizione del cliente attraverso differenti canali promozionali rende l'acquirente sempre più aggiornato e consapevole delle proprie scelte ma può ridurre i tempi di permanenza nel punto vendita. Risulta indispensabile pertanto saper gestire i big data generati dalle vendite e sfruttare le informazioni sugli acquisti e sul cliente per conoscerne meglio il comportamento e migliorare l'efficacia delle leve di marketing che influenzano le scelte del consumatore all'interno del punto vendita (quali il layout, il display merceologico etc.). Queste ultime, infatti, continuano ad essere strumenti importanti di differenziazione e fidelizzazione della clientela (Cardinali, 2009, p. 65). È dimostrato che il layout ha un effetto sulla frequenza di visita e sull'acquisto d'impulso ed è ancora determinante per le scelte d'acquisto in store (Mohan *et al.* 2013). Alla luce di questo panorama occorre fare leva sullo shopper marketing al fine di stimolare le decisioni di acquisto nel punto vendita (Shankar *et al.*, 2011) integrando questo sistema decisionale con le opportunità offerte dalle strategie di *online* e *mobile*

---

<sup>1</sup> Letizia Lo Presti, Dottore di ricerca in Marketing e Comunicazione presso l'Università degli Studi di Salerno, llopresti@unisa.it.

<sup>2</sup> Vittoria Marino, Professore associato in Marketing Internazionale, Università degli Studi di Salerno, vmarino@unisa.it

<sup>3</sup> Paolo Di Betta, Professore associato in Economia e Gestione delle Imprese, Università degli Studi di Palermo, paolo.dibetta@unipa.it

*marketing* (Shankar, 2014). Tecniche per l'elaborazione e la gestione manageriale dei dati come la *Market Basket Analysis* (MBA), che rileva le associazioni tra due o più prodotti, rendono tali dati informazioni utili da sfruttare per migliorare l'offerta nel punto vendita (*affinity analysis*). Ad oggi manca una connessione tra acquisti incrociati e layout, e negli studi che usano la MBA per l'analisi degli acquisti incrociati, mancano applicazioni concrete sul layout in grado di delineare modelli di acquisto ricorrenti. La nostra indagine tenta di rilevare i comportamenti di acquisto all'interno di uno *shopping experience* delineando i principali modelli comportamentali associativi (Contaldo e Largo, 2007) e individuando le principali categorie merceologiche che creano poli di attrazione all'interno del layout. Applicando la tecnica della MBA a 128.634 scontrini lo studio individua le *strong association rules* e rileva l'efficacia del layout nell'aumentare l'esperienza d'acquisto sul punto vendita.

### **Review della letteratura**

Lo studio del comportamento d'acquisto all'interno del punto vendita è determinante soprattutto in vista di una consolidata pratica che vede la programmazione degli acquisti realizzarsi sempre più *out-store* che *in-store* (Cardinali, 2009). Nell'era del 2.0 gli operatori di marketing devono considerare tutti gli strumenti *online* e *offline* che possono condizionare le scelte del consumatore nel lungo periodo (Shankar, 2014): dallo scaffale, al layout, alle informazioni in rete. Studi sul comportamento d'acquisto del consumatore nel punto vendita hanno per lo più applicato la tecnica del questionario e utilizzato variabili esogene (il layout, la musica e l'atmosfera) ed endogene (variabile sociodemografiche) per rilevare gli effetti sugli acquisti d'impulso (Mohan *et al.* 2013). Altri studi, seppur in numero limitato, hanno utilizzato la tecnica *Market Basket Analysis* (MBA) per rilevare i comportamenti d'acquisto ricorrenti dei clienti a partire dall'elaborazioni delle transazioni effettivamente realizzate nel punto vendita. La MBA, un'importante applicazione di marketing dei sistemi di *data mining* che individua referenze o gruppi di referenze che tendono ad essere acquistate assieme, è una metodologia che consente di studiare il comportamento di acquisto del cliente una volta chiusa la transazione nel punto vendita. Applicazioni recenti nel campo del marketing si sono concentrate principalmente sulle transazioni effettuate all'interno dei supermarket e hanno analizzato quali prodotti più frequentemente sono acquistati insieme in una stessa transazione (Kapadia and Kalyandurgmath, 2015; Mustafa, 2015; Hoanca and Mock, 2011). Tali studi sottolineano come la tecnica possa migliorare la formulazione del mix di marketing (in termini di prezzo e promozioni) nei supermarket (Mustafa, 2015), negli store di attrezzature sportive (Kaur e Singh, 2013) e nelle librerie universitarie (Hanca e Mock, 2011), migliorare i profitti, formulare strategie di *up-selling* e *cross-selling*, supportare attività di *customer relationship management* e incrementare l'efficacia del layout nel punto vendita. Nonostante questi studi abbiano il pregio di impiegare la MBA in contesti diversi dal tipico ambito di applicazione (il supermarket), essi adottano la tecnica su un campione ridotto di scontrini e sui prodotti acquistati piuttosto che sui livelli merceologici. Inoltre, anche se questi studi individuano come implicazioni manageriali la possibilità di utilizzare la MBA per riorganizzare il layout in modo da facilitare lo *shopping experience* dell'acquirente (Contaldo e Largo, 2007; Mostafa, 2015), non sembra ci sia ancora una reale

applicazione allo spazio espositivo tale da visualizzare il flusso del traffico a partire dai modelli comportamentali emersi dagli acquisti incrociati. In questo senso un recente studio ha applicato la MBA all'ambito dei trasporti pubblici (Sokolowska, 2014) analizzando e visualizzando l'andamento dei flussi dei passeggeri al fine di adattare le tratte dei bus alle reali esigenze dei cittadini. Sulla scia di questi studi, il presente lavoro cerca di visualizzare i comportamenti di acquisto incrociato ricorrenti e individua le principali regole associative che sul punto vendita creano maggior traffico.

### **Metodologia**

Per rilevare i reali comportamenti d'acquisto degli acquirenti nel punto vendita è stata effettuata un'analisi degli acquisti incrociati attraverso l'individuazione delle più ricorrenti regole associative su 33 categorie merceologiche. Tale analisi è stata condotta attraverso l'applicazione della tecnica della MBA. Per estrarre le regole associative del tipo  $A \rightarrow B$  ("nelle transazioni in cui compare A compare anche B") sono stati calcolati quattro indici: l'indice di *support* che misura la percentuale di transazioni che contengono entrambe le categorie merceologiche A e B; l'indice di *confidence* che misura quante volte gli elementi in B appaiono in transazioni che contengono A; l'indice *lift* che rappresenta una misura di normalizzazione utile a stabilire se tra due eventi vi sia una correlazione positiva o negativa: quando l'indice è  $>1$  la probabilità di acquistarli congiuntamente è superiore alle probabilità di acquistarne solo uno (Contaldo e Largo, 2007); l'*odds ratio* misura, per ogni coppia di prodotti, la forza dell'associazione tra le coppie di categorie merceologiche. Infine, sono state individuate le *strong association rules* a partire da un "minimo support" e una "minima confidence" (Giudici, 2005) pari o superiore, rispettivamente a 4% e 0,1. Sulla scorta di questa prima rilevazione sono state individuate le categorie merceologiche che calamitano di più gli acquisti ricorrenti e poi sono stati riportati sul layout del punto vendita per individuare le "aree calde" e le "aree fredde" (Lugli, 2009) dello spazio espositivo.

### *Campione*

Da un dataset composto da 715.662 scontrini è stato estratto un campione di 128.364 scontrini rilevati da un centro *shopping experience* sito nella città di Palermo. Per estrarre il campione oggetto dell'analisi è stato effettuato un campionamento a grappoli. Dai 304 giorni di effettiva apertura del punto vendita (eccetto le domeniche) sono stati estratti casualmente 72 giorni. Da ogni giorno così estratto è stato prelevato il totale degli scontrini per un ammontare complessivo di 184.441 transazioni. Questi poi sono stati depurati di tutti gli scontrini con un solo prodotto ottenendo un campione di 128.364 scontrini.

### **Risultati**

L'analisi degli scontrini-campione ha restituito 1056 regole associative. Sulla base del "minimo support" e della "minima confidence", solo 50 regole associative risultano significative. La tabella 1 illustra le *strong association rules* e i rispettivi indici. Delle 33 linee merceologiche solo 10 risultano essere coinvolte negli acquisti incrociati. Le categorie merceologiche coinvolte nelle regole associative sono Igiene per la casa/persona, Casalinghi, Arredamento, Alimentari, Arredo verde, Giocattoli,

Materiale elettrico, Illuminazione, Bruno, Ped. Esse sono alternativamente presenti o come testa o come corpo della regola in base al grado di associazione che le contraddistingue. Le categorie appartengono tutte al comparto della casa o della persona: mobili, televisori, piccoli elettrodomestici, igiene casa, lampadari, alimentari e giocattoli etc. che rispondono a bisogni tangibili e intangibili per lo più del cliente “famiglia”, cliente target del centro shopping experience oggetto dell’analisi.

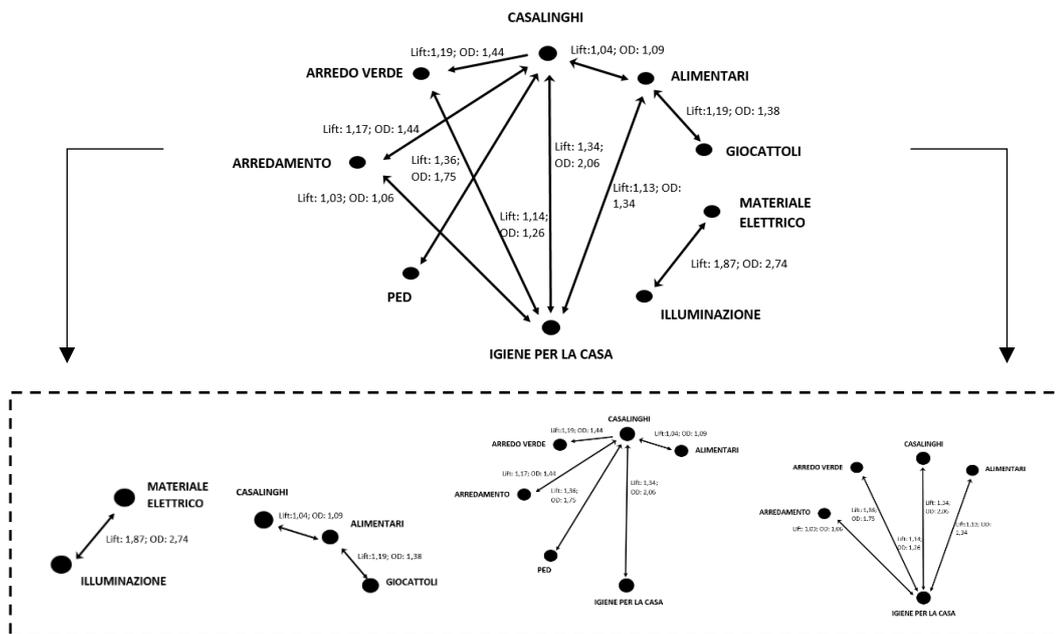
Tabella 1: Le *strong association rules*

REGOLA		Support (%)	Confidence	Lift	Odds ratio
Casalinghi=>	Igiene casa	14,14	0,45	1,34	2,06
Igiene casa=>	Casalinghi	14,14	0,42	1,34	2,06
Alimentari=>	Igiene casa	12,29	0,38	1,13	1,34
Igiene casa=>	Alimentari	12,29	0,37	1,13	1,34
Arredamento=>	Casalinghi	11,13	0,37	1,17	1,44
Casalinghi=>	Arredamento	11,13	0,35	1,17	1,44
Alimentari=>	Casalinghi	10,70	0,33	1,04	1,09
Casalinghi=>	Alimentari	10,70	0,34	1,04	1,09
Arredamento=>	Igiene casa	10,21	0,34	1,03	1,06
Igiene casa=>	Arredamento	10,21	0,31	1,03	1,06
Alimentari=>	Arredamento	9,17	0,28	0,94	0,89
Arredamento=>	Alimentari	9,17	0,31	0,94	0,89
Alimentari=>	Bruno	6,30	0,19	0,93	0,88
Bruno=>	Alimentari	6,30	0,30	0,93	0,88
Alimentari=>	Giocattoli	6,03	0,19	1,19	1,38
Giocattoli=>	Alimentari	6,03	0,39	1,19	1,38
Bruno=>	Igiene casa	5,09	0,24	0,73	0,58
Igiene casa=>	Bruno	5,09	0,15	0,73	0,58
Igiene casa=>	Materiale elettrico	4,96	0,15	0,96	0,94
Materiale elettrico=>	Igiene casa	4,96	0,32	0,96	0,94
Bruno=>	Casalinghi	4,92	0,24	0,75	0,61
Casalinghi=>	Articoli in stock	4,92	0,16	0,75	0,61
Igiene casa=>	Illuminazione	4,91	0,15	0,93	0,89
Illuminazione=>	Igiene casa	4,91	0,31	0,93	0,89
Casalinghi=>	Illuminazione	4,85	0,15	0,97	0,95
Illuminazione=>	Casalinghi	4,85	0,31	0,97	0,95
Giocattoli=>	Igiene casa	4,66	0,30	0,90	0,83
Igiene casa=>	Giocattoli	4,66	0,14	0,90	0,83
Illuminazione=>	Materiale elettrico	4,56	0,29	1,87	2,74
Materiale elettrico=>	Illuminazione	4,56	0,30	1,87	2,74
Casalinghi=>	Materiale elettrico	4,54	0,14	0,93	0,89
Materiale elettrico=>	Casalinghi	4,54	0,29	0,93	0,89
Casalinghi=>	Ped	4,52	0,14	1,36	1,75
Ped=>	Casalinghi	4,52	0,43	1,37	1,75
Arredamento=>	Illuminazione	4,48	0,15	0,95	0,92
Illuminazione=>	Arredamento	4,48	0,28	0,95	0,92
Alimentari=>	Materiale elettrico	4,48	0,14	0,89	0,83
Materiale elettrico=>	Alimentari	4,48	0,29	0,89	0,83
Arredamento=>	Bruno	4,29	0,14	0,69	0,55
Bruno=>	Arredamento	4,29	0,21	0,69	0,55
Casalinghi=>	Giocattoli	4,23	0,13	0,86	0,78
Giocattoli=>	Casalinghi	4,23	0,27	0,86	0,78
Alimentari=>	Illuminazione	4,21	0,13	0,82	0,72
Illuminazione=>	Alimentari	4,21	0,27	0,82	0,72
Arredo verde=>	Igiene casa	4,16	0,38	1,14	1,26
Igiene casa=>	Arredo verde	4,16	0,12	1,14	1,26
Arredo verde=>	Casalinghi	4,11	0,38	1,19	1,34
Casalinghi=>	Arredo verde	4,11	0,13	1,19	1,34
Arredamento=>	Giocattoli	4,09	0,14	0,88	0,81
Giocattoli=>	Arredamento	4,09	0,26	0,88	0,81

In particolare le categorie Casalinghi e Igiene per la casa sono quelle che calamitano l’acquisto di altre categorie essendo il corpo o la testa delle più ricorrenti regole associative. Infatti, come è possibile vedere dalla figura 1 che illustra i principali acquisti incrociati sulla base del calcolo dell’indice lift, emerge come siano principalmente 4 i *pattern* ricorrenti negli acquisti incrociati degli acquirenti. In tre

modelli su 4 le categorie Casalinghi e Igiene per la casa rappresentano i “poli d’attrazione” per gli acquisti della categoria Giocattoli, Alimentari, Arredamento, Arredo verde e PED. Ciò significa che acquistare Casalinghi e Igiene della casa aumenta la probabilità di avere anche una delle altre 7 categorie merceologiche. Un’altra forte associazione è quella tra la categoria Materiale elettrico e Illuminazione come dimostrato anche dall’odds ratio (OD: 2,74).

Figura 2: I modelli comportamentali associativi per categorie merceologiche



Collocando le *strong association rules* all’interno della mappa del punto vendita sulla base dell’indice “support” si può rilevare la presenza di una “area calda” (colore rosso) che crea maggiore traffico ed è collocata in buona parte della superficie espositiva e una “area fredda” (colore azzurro) che presenta un basso indice di passaggio. In particolare modo la figura 2 mostra un nodo centrale presso il quale sono convogliati tutti gli acquisti che coincide con le categorie Casalinghi e Igiene casa/persona. Nella “area calda” del punto vendita sono collocate le categorie merceologiche con più alta frequenza d’acquisto congiunto (Alimentari, Casalinghi, Igiene casa/persona, Arredamento) mentre nella “area fredda”, troviamo le linee merceologiche che presentano una bassa movimentazione e in cui risiedono per lo più i prodotti ad acquisto programmato quali frigoriferi, condizionatori, scaldabagni etc., tutti prodotti shopping good per i quali il consumatore, prima di decidere il prodotto da acquistare, impiega del tempo nel confrontare marche, qualità ed estetica.

Figura 2: Le *strong association rules* sul layout del punto vendita



### **Conclusioni e implicazioni manageriali**

I primi risultati di questo studio hanno evidenziato i principali modelli comportamentali dell'acquirente attraverso l'analisi delle categorie merceologiche che sono acquistate congiuntamente. Il lavoro, inoltre, applica le *strong association rules* alla mappa del punto vendita visualizzando le zone con più alta probabilità di acquisto congiunto e con basso indice di passaggio. La presenza di una area fredda nel punto vendita può comunque suggerire mirate azioni di shopper marketing che aumentino il traffico in-store. Pertanto, l'analisi della MBA sulle categorie merceologiche potrebbe aiutare gli store manager a valutare quali tra queste creano traffico nel punto vendita e determinare opportune strategie che possono massimizzare il margine complessivo per unità di spazio espositivo delle principali categorie merceologiche coinvolte nelle *strong association rules* migliorando l'efficacia del layout.

### **Bibliografia**

- Cardinali M.G. (2009). *Shopper Marketing. Creare valore nei luoghi*. Egea, Milano
- Di Giovanna Contaldo, R. e Largo, T. (2007). Strumenti di studio del portfolio prodotti: analisi associativa. *PMI* n. 10/2007, pp. 54-59
- Giudici P. (2005). *Data mining, metodi informativi, statistici e applicazioni*, McGraw Hill, Milano.
- Hoanca, B., & Mock, K. (2011). Using market basket analysis to estimate potential revenue increases for a small university bookstore. In *Conference for Information Systems Applied Research*, Vol. 4(1822) pp. 1-11.
- Kapadia, G., & Kalyandurgmath, K. (2015). Market Basket Analysis of Consumer Buying Behaviour of a Lifestyle Store. In *International Conference on Technology and Business Management*, March, Vol. 23, pp. 406-412.
- Kaur H. e Singh K. (2013). Market Basket Analysis of Sports Store using Association Rules, *International Journal of Recent Trends in Electrical & Electronics*, 3(1) pp. 81-85
- Lugli G. (2009). *Marketing distributivo*, Milanofiori Assago (Mi), Utet,
- Mohan, G., Sivakumaran, B., & Sharma, P. (2013). Impact of store environment on impulse buying behavior. *European Journal of Marketing*, 47(10), pp. 1711-1732.
- Mostafa, M. M. (2015). Knowledge discovery of hidden consumer purchase behaviour: a market basket analysis. *International Journal of Data Analysis Techniques and Strategies*, 7(4), 384-405.
- Shankar, V., Inman, J. J., Mantrala, M., Kelley, E., & Rizley, R. (2011). Innovations in shopper marketing: current insights and future research issues. *Journal of Retailing*, 87(1), pp. S29-S42.
- Sokołowska, D. (2014). Market Basket Analysis as a Support Tool for The Management of Public Transport. *Studies in Logic, Grammar and Rhetoric*, 37(1), pp. 219-237.