

**UNIVERSITÀ DEGLI STUDI DI PISA**  
Facoltà di Economia  
Facoltà di Scienze, Matematiche, Fisiche e Naturali  
Corso di Laurea specialistica in  
Informatica per l'Economia e per l'Azienda



**TESI DI LAUREA**

**VERSO LA BUSINESS INTELLIGENCE: IL PERCORSO DI  
UN'AZIENDA NEL SETTORE MODA**

RELATORE:

Dott.ssa Michela Ornati

RELATORE:

Prof.ssa Fosca Giannotti

RELATORE:

Prof. Nicola Ciaramella

Candidato:  
Mimosa Ronca

ANNO ACCADEMICO 2004-05



## **Ringraziamenti**

Ho completato il mio percorso universitario con un' esperienza di tirocinio in azienda che ha arricchito la mia formazione e mi ha dato modo di esprimere le capacità e conoscenze acquisite nelle aule universitarie. La tesi nasce sulla base di questa esperienza di lavoro e di formazione svolta durante il tirocinio, per cui devo ringraziare l'Azienda in cui ho svolto il lavoro, la Dott.ssa Michela Ornati, responsabile aziendale del tirocinio, e la Prof.ssa Fosca Giannotti, responsabile accademica.

In particolare esprimo tutta la mia riconoscenza a Michela per avermi insegnato a lavorare e per essere stata un esempio eccellente di professionista seria, capace e gentile. Ringrazio anche il mio collega Stefano per la sua amicizia e Alberto Massocchi per tutto quello che mi ha insegnato e per avermi sempre ascoltato pazientemente.

Per la scrittura e l'organizzazione della tesi mi rimane da ringraziare il mio relatore e professore Prof. Nicola Ciaramella.

Per la disponibilità e le attenzioni dedicate ai propri studenti ringrazio il presidente del corso di laurea specialistica in *Informatica per l'Economia e per l'Azienda* Prof. Antonio Albano.

Per quanto riguarda la mia esperienza universitaria da un punto di vista più personale non potrò mai ringraziare abbastanza i miei genitori, i cui consigli, indicazioni e soprattutto la cui stima mi hanno condotto felicemente ai traguardi raggiunti e faranno altrettanto per quelli che mi aspettano in futuro.

Ringrazio gli amici e i parenti che mi hanno incoraggiato e appoggiato e che mi vogliono bene, in particolare quelli che più da vicino hanno seguito il mio percorso di studi, mio fratello, mia nonna Iliana, Monica, Massimiliano, Maurizio, Alberto, Valeria, Oliviero, Tommaso, Michele, Nicolò e i miei compagni di corso.



1. Introduzione.....	7
2. Il <i>Customer Relationship Management</i> .....	11
1.1. L'impresa.....	14
1.2. Il valore del cliente.....	17
1.3. Aspettative dal CRM: <i>che cosa si aspettano in concreto i manager</i> .....	23
<i>quando decidono di applicare una strategia di CRM?</i>	
1.3.1. Alcuni esempi eccellenti.....	25
3. Un sistema di <i>business intelligence</i> per il CRM .....	29
3.1 Tecnologie a supporto CRM operativo.....	31
3.2 Tecnologie a supporto del CRM analitico .....	32
3.2.1 Il Datawarehouse.....	33
3.2.2 L'Analisi Multidimensionale.....	36
3.2.3 Il Data Mining .....	39
4. Sviluppo di un progetto di CRM.....	45
4.1 Contesto di Business.....	45
4.2 Il progetto di CRM .....	47
4.3 <i>Osservare che cosa il cliente sta facendo</i> .....	50
4.3.1 Azioni da intraprendere.....	53
4.4 <i>Ricordare che cosa i clienti hanno fatto nel passato</i> .....	54
4.4.1 Azioni da intraprendere.....	57
4.5 <i>Imparare da ciò che ci ricordiamo</i> .....	58
4.5.1 L'analisi dei dati.....	60
4.5.2 Strumento di analisi.....	63
4.5.3 Analisi per la gestione delle campagne marketing.....	65
4.6 <i>Comportarsi conseguentemente a ciò che si è imparato per rendere</i> <i>i clienti più profittevoli</i> .....	78

5. Idee per estendere l'ambiente di <i>business intelligence</i> .....	83
5.1 L'ambiente di <i>business intelligence</i> attuale.....	83
5.2 L'ambiente di <i>business intelligence</i> verso cui si vuole andare.....	86
5.3 Analisi dei requisiti di un ambiente di <i>business intelligence</i> .....	90
5.4 Fasi dell'implementazione.....	106
6. Conclusioni.....	122
APPENDICE.....	125
Bibliografia.....	131
Bibliografia in rete.....	132

## 1. Introduzione

Nel corso di questa trattazione si discuteranno i temi che riguardano le problematiche incontrate da un'azienda nell'affrontare per la prima volta lo sviluppo di un ambiente di *business intelligence*, analizzando sia gli aspetti aziendali ed organizzativi sia quelli informatici e tecnologici. Gli argomenti trattati verranno approfonditi attraverso la descrizione dell'esperienza di un'Azienda italiana che opera nel settore moda, in cui l'autore ha seguito in prima persona e attivamente buona parte delle fasi dello sviluppo dell'ambiente di business intelligence, potendo così osservare come questo ambiente e l'organizzazione evolvano in maniera incrementale ed interattiva.

La peculiarità del contenuto di questo lavoro è che l'esperienza realmente vissuta dall'Azienda viene raccontata e descritta approfonditamente attraverso lo svolgersi delle sue fasi critiche scandite da problematiche e opportunità e relative soluzioni.

La letteratura specializzata è infatti caratterizzata, nel campo aziendale, da un approccio troppo generale o troppo teorico difficilmente ricollegabile ai problemi che si incontrano nelle situazioni reali e, nel contesto informatico, è diffusa una letteratura specialistica che affronta temi tecnologici e scientifici spesso molto distanti dai contesti di applicazione e inaccessibili ai non addetti ai lavori. Inoltre è opinione dell'autore che un ambiente di business intelligence, per la complessità della sua natura e del suo contesto di applicazione, non può che essere descritto se non tenendo conto sia degli aspetti organizzativi e strategici sia degli aspetti tecnologici presenti nell'azienda.

L'Azienda del caso di studio preso in considerazione si è avvicinato alla business intelligence attraverso il percorso di un progetto di *customer relationship management*, per cui la trattazione si focalizza principalmente sull'ambiente e sul sistema di *intelligence* applicato per il supporto della strategia di CRM.

Il CRM è una strategia che mette al centro dell'azienda il cliente e il suo valore; si basa sulla costruzione di una relazione con il cliente che procuri la sua soddisfazione e lo porti verso un comportamento fedele e di *loyalty* verso l'azienda. La relazione deve essere strutturata in maniera tale che l'azienda sia in grado di ascoltare e osservare i comportamenti del cliente per accumulare una conoscenza su di essi. Sulla base di questa conoscenza costruire quindi i presupposti per una relazione mirata alla

soddisfazione di ogni individuo che interagisce con l'impresa. Il CRM si basa sul ciclo di monitoraggio e raccolta delle informazioni, analisi dei dati, azione e misurazione dei risultati e pertanto presuppone l'esistenza di un'infrastruttura informatica per la raccolta e la gestione dei dati e di un sistema di intelligence per l'analisi e la produzione di informazioni.

La necessità di un ricorso massiccio a sofisticate e costose tecnologie informatiche ha posto un freno alla diffusione di tale strategia nella maggior parte delle imprese italiane, mentre non ha ostacolato quelle aziende, soprattutto statunitensi e inglesi, che già da alcuni decenni hanno potuto e voluto disporre delle tecnologie informatiche più potenti e all'avanguardia.

Il caso aziendale oggetto di studio nella tesi è emblematico di un movimento e di una tendenza in corso in questi ultimi anni nelle aziende italiane verso l'applicazione di strategie di CRM e l'adozione di sistemi di business intelligence. Questo trend ha coinvolto anche le aziende più piccole e quelle che fino a pochi anni fa, per cultura aziendale e per tipologia di business, non avevano incluso tra le proprie competenze distintive quelle informatiche.

Come emergerà dall'analisi dei sistemi di CRM e del ruolo che in tali sistemi hanno le tecnologie di *business intelligence*, la *business intelligence* è considerata molto di più di un insieme di tecnologie che possono, se opportunamente utilizzate ed implementate, creare valore per un'impresa ma è ritenuta essa stessa un valore, un vantaggio competitivo che nelle condizioni economiche attuali è in grado di giocare un ruolo determinante per la sopravvivenza dell'impresa.

Nel capitolo 2 *Il Customer Relationship Management*, viene fatta una presentazione della strategia di CRM, discutendo in particolare delle motivazioni che spingono tante aziende verso un approccio cliente-centrico e di quali sono i maggiori benefici che si ricavano dall'applicazione di una tale strategia.

Si dà un inquadramento storico ed economico delle strategie di CRM e si individua nella teoria aziendale una tipologia di impresa che si adatta al contesto economico attuale, l'*impresa proattiva*.



L'*impresa proattiva* e cliente-centrica è alla ricerca costante della comprensione del cliente per essere in grado di gestire la relazione con esso nel modo migliore possibile e massimizzando il valore che il cliente ha per l'impresa stessa.

Dopo aver fornito l'inquadramento teorico necessario si torna al nocciolo del problema: che cosa significa e quali sono i vantaggi di una strategia di CRM che i manager cercano; dal momento che in alcuni settori l'approccio al CRM è ampiamente diffuso e collaudato, il manager che affronta il problema per la prima volta può trarre ispirazione sia dai molteplici esempi riportati dalla stampa specializzata sia dall'esperienza vissuta in prima persona come consumatori.

Nel capitolo 3 *Un sistema di business intelligence per il CRM*, si fornisce una panoramica sulle tecnologie che comunemente costituiscono l'ambiente di *business intelligence* a supporto del CRM.

Vengono identificati due famiglie distinte di applicazioni informatiche per il CRM, quelle per il *CRM operativo* e quelle per il *CRM analitico*. Queste due famiglie di applicazioni sono distinte ma negli ambienti di *business intelligence* convivono e interagendo formano l'ambiente stesso.

Le tecnologie di CRM operativo sono quelle che servono per automatizzare il reperimento delle informazioni e la memorizzazione dei dati provenienti dal cliente, sono tecnologie di DBMS per la gestione ordinata ed efficiente delle basi di dati e applicazioni per la gestione delle interfacce di interazione, come il sito web, le casse in negozio, i call center, ecc.

Le tecnologie per il CRM analitico sono rappresentate dal *datawarehousing* e dal data mining. Il datawarehouse è un archivio di dati che memorizza e organizza i dati perché possano essere analizzati per produrre indicatori di sintesi a livello direzionale.

Le analisi che si possono fare sul datawarehouse, e che non si possono fare agevolmente su un database operativo, sono di tipo multidimensionale, permettono cioè di osservare un fenomeno complesso da più punti di vista e a un livello di sintesi molto superiore a quello della transazione (livello di dettaglio a cui si interrogano i database operativi).

Il data mining è una tecnologia che permette di estrarre conoscenza non nota a priori da grandi masse di dati, la produzione di tale conoscenza è il momento conclusivo di un

processo, il *Knowledge Discovery in Databases*, che coinvolge gli esperti del business, dei dati e del data mining.

Nel capitolo 4 *Sviluppo di un progetto di CRM*, si presenta l'Azienda che costituisce il caso di studio e vengono descritte le caratteristiche del progetto di CRM che l'Azienda ha intrapreso.

Si analizzano nel dettaglio le motivazioni che hanno spinto l'Azienda ad intraprendere tale progetto e le modalità con cui lo ha condotto.

Vengono descritte le problematiche e le soluzioni trovate compatibilmente con i vincoli organizzativi e tecnologici, e le azioni che andrebbero intraprese per migliorare il processo o correggerlo. Il processo si divide in quattro fasi:

1. Osservare che cosa il cliente sta facendo
2. Ricordare che cosa i clienti hanno fatto nel passato
3. Imparare da ciò che ci ricordiamo
4. Comportarsi conseguentemente a ciò che si è imparato per rendere i clienti più profittevoli

Nel capitolo 5 *Idee per estendere l'ambiente di business intelligence*, si discute della nuova fase di vita del progetto di CRM e delle nuove opportunità che si sono aperte sulla base dei primi risultati ottenuti.

Si analizza l'eventualità di estendere l'ambiente di *business intelligence* attuale con un sistema di data mining conducendo uno studio di fattibilità che oltre ad analizzare le risorse disponibili propone delle possibili soluzioni per ampliare l'ambiente e migliorarlo.

Infine, utilizzando il modello standard di processo CRISP-DM, viene descritto come sarebbe un processo di data mining con il nuovo ambiente di *business intelligence*.

## ***2. Il Customer Relationship Management***

Il CRM è una strategia di business che mira a identificare, attrarre e mantenere i migliori clienti al fine di massimizzare la profittabilità delle relazioni con il cliente.

Lo scopo del CRM è sviluppare relazioni personali con i propri clienti, conoscerli e soddisfarli per renderli fedeli all'azienda o ad un prodotto.

Nei decenni passati per far fronte alla forte competizione si sono adottate principalmente due strategie, la politica di efficienza dei costi e la strategia di apprendimento finalizzata alla proposta sul mercato del prodotto più innovativo o qualitativamente migliore ma nella maggior parte dei contesti competitivi attuali entrambe queste strategie non sono sufficienti per portare l'azienda al successo di mercato. Ciò che può dare un vantaggio competitivo ai giorni nostri è la capacità di soddisfare il cliente a tal punto che reiteri l'acquisto e rimanga fedele alla marca.

Per la maggior parte dei settori industriali e commerciali il contesto storico attuale si caratterizza per avere un carattere globale che se da una parte ha aperto nuovi mercati di sbocco (consumatori finali) dall'altra ha acuito la competizione. Per aziende che producono nel territorio americano o europeo mettere in atto strategie di leadership di costo è diventato ormai molto difficile e comunque una leadership basata sul prezzo può essere solo temporanea vista l'accessibilità in termini di costo alla tecnologia.

Il consumatore attuale sta imparando a muoversi anch'esso in un contesto globale (internet e mezzi di comunicazione tradizionali), ha a disposizione una grande quantità di informazioni e proposte diverse e pertanto le sue aspettative sono sempre più alte e mirate. Un consumatore che si muove liberamente e agevolmente in questo contesto sarà sempre in cerca della migliore soluzione alle proprie esigenze e la chiave del successo competitivo è fare in modo di ritagliargli su misura la prossima offerta in maniera tale che il cliente non abbia nessun motivo per rivolgersi alla concorrenza.

Perché il cliente sia fedele ad un'azienda deve essere soddisfatto dall'intera esperienza che ha intrattenuto con questa, dove l'esperienza comprende l'intero ciclo che va dal contatto indiretto attraverso la pubblicità, al contatto diretto durante l'acquisto e il consumo ed infine all'esperienza post-acquisto o consumo. Lo scopo del CRM è pertanto quello di costruire una relazione duratura con il cliente, relazione che ha un

duplice uso, è un canale attraverso cui reperire informazioni sul cliente e rappresenta l'immagine stessa con cui l'azienda si rapporta al proprio cliente.

La relazione tra azienda e cliente idealmente viene identificata come uno-a-uno. L'azienda e suoi manager dovrebbero avere la stessa conoscenza del cliente che ha il personale di vendita che intrattiene relazioni dirette, personali e continuative con la propria clientela; data l'impossibilità di stabilire *relazioni umane* uno-a-uno tra l'azienda e ogni suo cliente l'obiettivo è di rendere condivisibile e condivisa la conoscenza che emerge da ogni singolo contatto avvenuto (in negozio, sul sito web, call center, ecc.).

Il concetto di relazione basato sulla conoscenza è senza dubbio ben noto alle imprese di servizi che solitamente offrono un servizio totalmente personalizzato e pertanto mantengono un contatto diretto con il cliente per tutta la durata del processo. Le imprese produttive o commerciali che si avvicinano al CRM dovrebbero pensare a loro stesse come ad un'impresa di servizi; esistono numerosi casi in cui l'impresa si è effettivamente attrezzata a diventare un'erogatrice di servizi oltre che di prodotti, si pensi ad esempio all'importanza che si attribuisce alla presenza e alla qualità dell'assistenza post-vendita nelle decisioni di acquisto di un PC o un elettrodomestico.

Per conoscere il cliente pertanto le intuizioni dei manager non bastano, bisogna far parlare i dati raccolti durante le transazioni effettuate dal cliente e trasformare tali dati in informazioni fruibili che possano essere di supporto e guida per prendere le decisioni operative e strategiche.

La conoscenza del cliente non può essere patrimonio esclusivo del Marketing, il CRM è una strategia che attraversa e coinvolge tutte le aree aziendali, così come le informazioni sul cliente infatti possono arrivare da qualsiasi punto di contatto tra il cliente e l'azienda, qualsiasi area aziendale o processo deve tendere alla comprensione e alla soddisfazione dei propri clienti.

L'attenzione quindi è sul cliente e non solo sul prodotto o sulla vendita come è accaduto comunemente negli ultimi decenni. Per convertire un'organizzazione orientata al prodotto o alla vendita in un'organizzazione cliente-centrica la strategia di CRM coinvolge l'azienda nella sua globalità e deve essere supportata da un'infrastruttura tecnologica molto sofisticata. Si tratta quindi di una strategia di lungo termine che presuppone ingenti investimenti con ritorni difficilmente quantificabili nel breve.

Il CRM si contraddistingue per avere un approccio quantitativo, la strategia per la creazione di una relazione con il cliente si applica nelle seguenti attività:<sup>1</sup>

- *Osservare che cosa il cliente sta facendo*; lo scopo è quello di accentrare tutta la conoscenza esistente sul cliente facendo confluire tutte i dati e le informazioni presenti nelle diverse aree aziendali e passanti dai vari canali. “*Il sistema di CRM deve essere gli occhi e gli orecchi dell’azienda e forse anche il naso, la lingua e le dita*” [Berry, Linoff 1997]
- *Ricordare che cosa i clienti hanno fatto nel passato*; essere cioè in grado di memorizzare le informazioni che si accumulano nel tempo e di organizzarle secondo una logica che mantenga il punto di vista sul cliente.
- *Imparare da ciò che ci ricordiamo*; elaborare tali dati per trasformarli in informazioni comprensibili e utili per capire il cliente.
- *Comportarsi conseguentemente a ciò che si è imparato per rendere i clienti più profittevoli*; una volta estrapolate le informazioni più rilevanti prendere delle decisioni in base a queste e agire con lo scopo di rendere i clienti più profittevoli (ad esempio, attività per incentivare il *cross selling* o per mantenere i clienti a rischio di abbandono)

L’importanza della disponibilità dei dati ha fatto sì che quei settori aziendali che da lungo tempo hanno automatizzato le transazioni con i propri clienti sviluppassero più facilmente e velocemente sistemi di CRM, settori quali ad esempio le banche, le assicurazioni e la grande distribuzione. Alcune delle aziende che operano in questi settori sono diventate delle vere e proprie venditrici di dati e informazioni sui clienti, la maggior parte delle compagnie di carte di credito per esempio vendono spazio pubblicitario sul foglio di comunicazione dei pagamenti inviato periodicamente al cliente; il prezzo che la compagnia di carte di credito può far pagare per lo spazio è proporzionale a quanto accuratamente è in grado di identificare il cliente che risponderà positivamente alla pubblicità.

Parallelamente alla grande disponibilità di dati il cambiamento che più ha dato la spinta alla diffusione dei sistemi di CRM è la drastica riduzione dei costi della tecnologia contemporaneamente all’aumento della potenza di calcolo.

---

<sup>1</sup> Si veda M.J.A Berry e G.S. Linoff “Data Mining Techniques For Marketing, Sales, and Customer Relationship Management” Second Edition, Wiley Publishing, Inc.

I progressi tecnologici degli ultimissimi anni, in particolar modo la diffusione dell'uso di Internet per le transazioni economiche, l'aumento di banda di trasmissione a basso costo e l'aumento della potenza di calcolo, hanno contribuito anche nelle situazioni più piccole, ma non solo, ad aumentare l'applicazione di strategie di CRM,

Per quanto riguarda la situazione italiana solo negli ultimissimi anni si stanno sperimentando i primi sistemi di CRM per due motivi principali, perché il panorama economico è costituito principalmente da piccole e medie imprese che avevano difficoltà a sostenere gli investimenti ingenti necessari per l'implementazione di un sistema di CRM e perché la maggior parte dei settori industriali è stata caratterizzata da una grande attenzione al prodotto e quindi spesso da un'organizzazione centrata sul prodotto o sulla vendita, come ad esempio i settori della moda e del lusso.

## **2.1 L'impresa**

La strategia di CRM si inserisce in un contesto teorico più ampio che individua e formalizza una nuova modalità di essere dell'impresa che si adatta alle caratteristiche della competizione globale dei nostri giorni: l'impresa proattiva<sup>2</sup>.

L'impresa proattiva è l'impresa che agisce in un contesto ambientale fortemente competitivo e sempre in cambiamento e che pertanto è proiettata verso il futuro, l'innovazione e la creatività.

Essere proiettati verso il futuro significa essere in grado di gestire eventi e conseguenze di eventi che non si sono ancora manifestati, di progettare il proprio futuro mentre si governa il presente e di definire le proprie mosse competitive sulla base della conoscenza del settore, di come è oggi e come sarà domani.

Per essere proiettati nel futuro si deve essere supportati dall'innovazione e si deve produrre innovazione stimolando e incoraggiando la creatività e lo spirito imprenditoriale di ogni componente dell'impresa.

L'impresa proattiva è tesa ad ascoltare sempre con la massima attenzione la voce dei *clienti* per soddisfarli al meglio; a questo fine essa deve rilevare le loro esigenze e preferenze presenti e anticipare quelle future, soddisfare e coltivare quelle attuali e suscitare e dirigere quelle potenziali. In questo suo sforzo, l'impresa deve tenere conto

---

<sup>2</sup> Si veda E. Valdani "L'impresa proattiva", 2000, McGraw-Hill.

della *concorrenza*, per comprendere e anticiparne le mosse e le tendenze evolutive. Per riuscire l'impresa deve sviluppare *capacità* distintive che le diano vantaggi competitivi e le consentano di dominare i fattori critici di successo nel suo mercato. Deve essere orientata al *cambiamento*, che nell'epoca attuale è ormai un dato acquisito del gioco competitivo, accettando come normale e vitale l'innovazione di idee, capacità, comportamenti, modelli organizzativi, ruoli delle persone, valori. Deve cercare il *consenso* di tutti gli agenti rilevanti, dal personale alle istituzioni alla collettività sociale. Deve avere un forte e deciso *convincimento* a perseguire la missione che si dà.

In sintesi, la formula imprenditoriale appena descritta si basa su:

- Clienti
- Concorrenza
- Capacità
- Cambiamento
- Consenso
- Convincimento

Per realizzare tale formula imprenditoriale l'impresa deve mantenere costantemente l'adesione con il mercato e per far questo deve coltivare lo sviluppo delle proprie capacità di osservazione, reazione ed anticipazione delle esigenze espresse dal mercato per sfruttare la conoscenza che ne deriva e crearsi nuovi vantaggi competitivi. Un'impresa che sviluppa tali capacità è in grado di mettere in atto economie di varietà (o personalizzazione), velocità, qualità e di apprendimento in base a ciò che il mercato chiede.

In particolare in questo contesto è interessante la realizzazione di un'economia di apprendimento ovvero la capacità di creare conoscenza e di utilizzarla attraverso:

- la sperimentazione continua di nuove idee e modalità di comportamento;
- l'apprendimento dall'esperienza;
- l'apprendimento da ciò che accade all'esterno dell'impresa;
- il trasferimento tempestivo ed efficace della conoscenza appresa a chi deve decidere ed agire.

L'impresa proattiva è un sistema adattivo complesso basato sull'apprendimento, l'apprendimento a sua volta si basa sul ciclo che comprende le capacità di raccogliere, coltivare, trasferire e applicare la conoscenza.

L'impresa proattiva viene rappresentata con la metafora molto incisiva e suggestiva del tempio, riportata in figura 1.

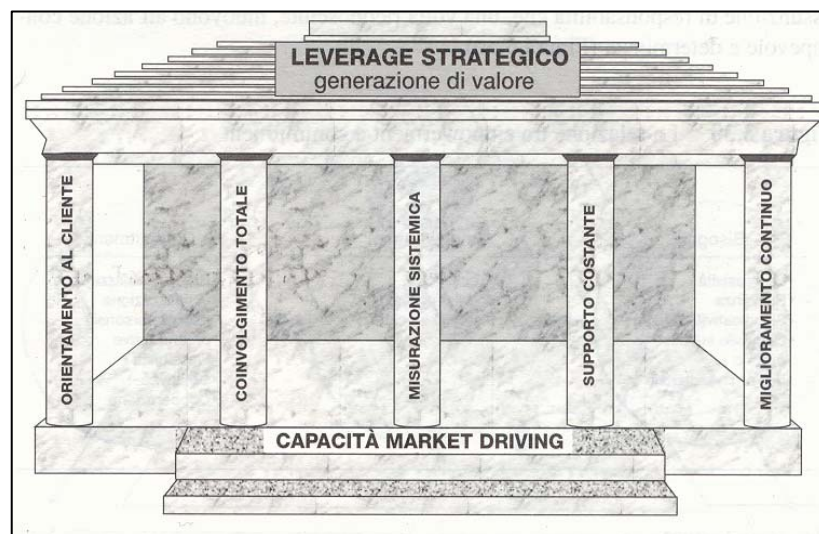


Figura 1: Il tempio e i pilastri dell'impresa proattiva [Fonte: Valdani 2000]

Le *capacità market driving*, le fondamenta, sono quelle descritte precedentemente che consentono di mettere in atto economie di varietà, velocità, qualità e apprendimento; permettono di adattare continuamente l'impresa ai cambiamenti già in atto nel suo ambiente e di provocare i cambiamenti stessi per creare un ambiente più favorevole.

L'*orientamento al cliente*, il primo pilastro, richiede che in ogni sua azione l'impresa tenga sempre fissata l'attenzione sulla voce del cliente (le sue preferenze e aspettative), sulla migliore proposizione di valore che si può offrire al cliente, sulla soddisfazione che il cliente ha avuto dalle transizioni passate e che potrà avere da quelle future con l'impresa.

Il pilastro del *coinvolgimento totale* indica la capacità dell'impresa di creare coesione intorno alle strategie ed obiettivi, attraverso la condivisione delle informazioni e la partecipazione delle persone, a tutti i livelli, nel processo decisionale.

Il pilastro della *misurazione sistematica* indica che l'impresa proattiva misura i fenomeni che la riguardano, sia interni che esterni, ogni volta che questo sia possibile ed



utile. Le misurazioni devono essere fatte per produrre informazioni sia conoscitive sia direzionali e motivazionali riassunte e sistematizzate in un cruscotto di indicatori sintetici ed espressivi dei vari aspetti dell'azienda.

Il pilastro *supporto costante* significa che l'impresa consuma e rigenera le proprie risorse pensando al futuro, programmando cioè uno sviluppo sostenibile nel tempo.

L'ultimo pilastro, il *miglioramento continuo* si riferisce all'uso che l'impresa proattiva fa dell'innovazione. L'innovazione è sistematica, frutto di un ciclo continuo di ideazione, misurazione, valutazione ed apprendimento che si traduce di nuovo in ideazione e così via.

Il tetto del tempio rappresenta la generazione di valore per il cliente e il *leverage* strategico.

Per massimizzare la generazione del valore si dovrà quindi individuare i bisogni, progettare e realizzare prodotti atti a soddisfarli e contenere i costi di produzione a un livello giustificato dall'importanza del prodotto per il cliente.

Il leverage strategico è il frutto delle capacità market driving e dei pilastri del valore, ed esprime la capacità dinamica dell'impresa di generare valore mantenendosi aperte le opzioni che le consentiranno di generarne ancora in futuro.

### **2.3 Il valore del cliente**

In questo contesto ciò che ci interessa di più è il primo pilastro su cui poggia l'impresa proattiva, l'orientamento al cliente, che indica il fatto che l'attenzione sia sui clienti, l'impresa infatti esiste per servire i propri clienti e per farlo al meglio deve essere capace di sapere come sono e cosa vogliono ora, prevedere come saranno e cosa vorranno in futuro, capire come agire per soddisfarli meglio dei competitori.

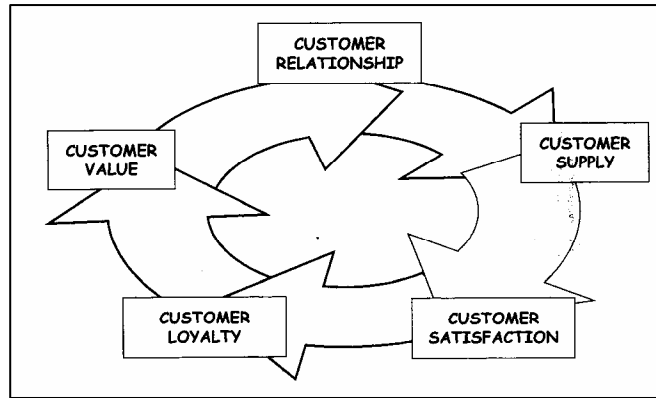
Il motivo per cui è così importante conoscere e capire i propri clienti è che l'impresa cliente centrica è l'impresa che considera la clientela la fonte primaria di generazione del valore. Dunque, il patrimonio fondamentale dell'impresa è la sua clientela. Tutti gli altri *asset* sono strumentali alla creazione e allo sviluppo delle relazioni con i clienti: sono queste relazioni che generano valore.

La creazione del valore è determinata in maniera preponderante dalle relazioni con la clientela. In questo caso è necessario approfondire che cosa si intende per relazione e

come l'impresa deve gestire tale relazione per sfruttare il suo valore in termini di profitto.

In termini generici per relazione si intende il ricordo delle esperienze passate, è cioè dal valore percepito dal cliente nelle passate interazioni con l'azienda che inizia il ciclo della redditività del cliente.

1. Dalla combinazione delle sue risorse e capacità l'impresa mette in atto il processo di monitoraggio, analisi da cui emerge una proposta basata sulla conoscenza del cliente. Tale proposta è una proposizione di valore per il cliente e che a seconda dell'efficienza dei processi e dell'efficace utilizzo delle competenze e conoscenze presenti nell'azienda si traduce in un'offerta di successo o meno.
2. La soddisfazione dei clienti, determinata dal raggiungimento delle aspettative, è funzione della proposizione di valore offerta dall'impresa.
3. La delizia è quello stato d'animo che si verifica quando le aspettative sono state superate e si riscontrano, per l'azienda, punti di forza da massimizzare.
4. La fedeltà dei clienti è funzione della soddisfazione e della delizia dei clienti. In termini pratici la *customer loyalty* si traduce in ripetizione dell'acquisto da parte del cliente che è rimasto soddisfatto degli acquisti precedenti.
5. Il valore delle relazioni con i clienti è funzione della fedeltà dei clienti, cioè è naturale assumere che la fedeltà dei clienti generi valore.
6. Il valore dell'impresa è funzione del valore delle sue relazioni con i clienti.
7. Il ciclo si chiude e si riapre con l'assunto che le risorse, le capacità e i processi dell'impresa sono funzione del valore dell'impresa. Ciò significa che l'esperienza acquisita durante il ciclo di proposizione dell'offerta, generazione di valore per il cliente, incontro della sua soddisfazione, incoraggiamento alla fedeltà, generazione di valore per l'impresa possano apportare miglioramenti e maggiori informazioni a sostegno del *customer relationship*. E che soprattutto i risultati ottenuti ed emersi durante il ciclo della redditività del cliente incoraggino il management verso nuovi investimenti per migliorare i processi, aumentare la conoscenza e migliorare l'offerta e la relazione.



**Figura 2: la catena della redditività della gestione del cliente**

Occorre però chiarire ancora alcuni punti distintivi del ciclo appena descritto ed in particolare i legami funzionali:

1. *Come l'impresa genera valore per il cliente attraverso la relazione.*

L'impresa cliente centrica crea valore mettendo in campo le proprie capacità di: 1) acquisire nuovi clienti, 2) sviluppare la redditività dei clienti in portafoglio, 3) fidelizzare e quindi mantenere i clienti. I tipici problemi da affrontare sono: a) come le scelte e le preferenze dei clienti influenzano la possibilità di generare valore, b) come la proposizione di valore debba essere progettata e offerta per incrementare il valore della relazione con la clientela, c) come il valore risultante dalla relazione debba essere suddiviso fra acquirente e venditore.

Il valore delle relazioni dipende a sua volta da ampiezza, intensità e qualità della rete di relazioni instaurate con i clienti.

2. *Come viene percepito il valore da parte del cliente*

Il cliente percepisce il valore della relazione in maniera diversa a seconda della fase della relazione in cui si trova e il valore percepito a sua volta spinge la relazione verso fasi diverse.

- *valore atteso*: il rapporto tra i benefici che ci si attendono dall'acquisto e dal godimento del prodotto e i costi che si ritiene di dover sostenere, anche in relazione alle alternative disponibili;

- *valore percepito dopo l'acquisto e l'uso*: deriva dalla soddisfazione o insoddisfazione e viene confrontato con il valore atteso;
- *valore percepito dopo le prime esperienze d'uso*: in termini comparativi rispetto alle alternative considerate; nasce dalla riflessione interna all'individuo cliente, che compara le esperienze certe maturate nella relazione con l'impresa con le aspettative ipotetiche verso le alternative;
- *valore di equità*: rapporto tra il valore che il cliente ritiene di avere ottenuto e quello che ritiene di aver generato per l'impresa.

### 3. *Come dal valore si genera Soddisfazione.*

La soddisfazione dipende dall'adeguatezza della proposizione di valore, che dipende dalle capacità e risorse dell'impresa e dai processi che utilizzano queste capacità e risorse. Ma risorse e capacità devono essere create e continuamente potenziate dagli investimenti degli azionisti, i quali sono motivati a farlo dalla crescita attuale e potenziale del valore dell'impresa. La soddisfazione è un atteggiamento e un sentimento molto complesso che sappiamo essere la mediazione tra le aspettative del cliente e l'effettiva esperienza di consumo e di interazione con l'azienda. Non si può sapere a priori quanto la proposizione di valore influisca sulla soddisfazione del cliente, o quanto la fedeltà influisca sulla redditività delle relazioni, questa sulla motivazione degli azionisti a investire nel potenziamento della capacità dell'impresa, e quanto queste influiscano sul miglioramento della proposizione di valore e quindi sulla soddisfazione della clientela.

### 4. *Come dalla Soddisfazione si arriva alla Fedeltà*

È intuitivo che la soddisfazione del cliente ne aumenti la fedeltà, ma non si può sapere a priori di quanto. In mercati dove esiste una molteplicità di offerte in competizione tra di loro il cliente tende costantemente a cercare alternative, e quindi un piccolo incremento di soddisfazione genera un piccolo incremento di fedeltà del cliente cosa che non succede invece in mercati regolamentati in presenza di situazioni semi o totalmente monopolistiche. Si hanno quindi gradi

diversi di intensità del legame soddisfazione-fedeltà, essendo diversa di caso in caso l'elasticità della risposta del cliente. Tra la soddisfazione e la fedeltà esiste lo stadio intermedio della delizia che, se raggiunto, apre la strada in maniera molto positiva alla prosecuzione della relazione verso la *loyalty*.

#### 5. *Quanto la fedeltà genera valore per l'impresa*

Come detto per i due punti precedenti è difficile stabilire a priori l'intensità del legame o la regola che collega il comportamento fedele di un cliente alla sua redditività e profittabilità per l'azienda.

Come emerge dalle considerazioni fatte sul modello teorico del ciclo della redditività del cliente c'è la necessità che i concetti di natura psicologica e comportamentale come soddisfazione e fedeltà debbano essere formalizzati e resi misurabili per poter arrivare alla produzione di informazioni che servano per prendere decisioni. C'è l'esigenza di modelli formali che esprimano cosa è una relazione, quali sono i fattori che la fanno nascere ed evolvere, come si misura il valore di una relazione. Le fasi della relazione descritte di seguito mettono in evidenza come il valore viene percepito dal cliente diversamente nel corso della relazione e come questo influenzi il corso e il valore stesso della relazione.

1. *Fase della soddisfazione e della accumulazione di fiducia*: la scelta d'acquisto è fondata sulla percezione da parte del cliente che sia la scelta migliore che può fare rispetto alle alternative. La percezione di soddisfazione è determinata in base alla differenza tra il valore atteso e quello percepito.
2. *Fase della fedeltà comportamentale*: le esperienze maturate che hanno generato un certo grado di soddisfazione alimentano la tendenza al riacquisto e progressivamente portano la relazione verso la fedeltà.
3. *Fase della fedeltà mentale*: la fedeltà mentale è la convinzione del cliente che l'impresa sarà capace di mantenere nel tempo un'offerta superiore rispetto alle alternative che esistono sul mercato. In questo stadio il cliente è pronto e disponibile per aumentare il valore della propria relazione. La sua attenzione verso le offerte della concorrenza si riduce e può anche essere disposto a mantenersi fedele

nonostante l'emergere di influenze contrastanti, come nuovi evidenti vantaggi nell'offerta di concorrenti.

4. *Fase della lealtà*: in questa fase il cliente sposta l'attenzione verso il valore di equità e inizia a considerare il valore che gli porta l'impresa in proporzione al valore portato dal esso all'impresa. Se il cliente pensa che lo scambio sia equo arriva alla lealtà.

Questo quadro teorico è teso a mettere il cliente e la relazione con il cliente al centro del processo conoscitivo, valutativo e decisionale. Gli assunti alla base del paradigma dell'impresa cliente centrica non sono affatto così scontati come si potrebbe essere indotti a pensare.

Di fatto si cerca di controllare una variabile esogena che per la maggior parte delle aziende è sempre rimasta fuori dalla porta. Il tentativo è di andare oltre al controllo e di aprire le proprie porte, le proprie capacità e il proprio modo di lavorare a ciò che di più complesso e per certi aspetti imprevedibile c'è, l'essere umano.

Che la soddisfazione del cliente migliori il suo atteggiamento è esperienza comune, ma un sistema di business intelligence, per essere utile come supporto alle decisioni, deve poter indicare *come* la soddisfazione genera l'atteggiamento positivo, e soprattutto *quanto* contribuisce a generarlo, e quanto invece l'atteggiamento positivo dipende da altri. La business Intelligence raccoglie una serie di tecnologie per supportare e per dar vita ad un processo in grado di rispondere a domande molto complesse in maniera decisiva.

Nel prossimo capitolo viene data una descrizione delle tecnologie della *business intelligence* applicate per la strategia di CRM.

### **1.3 Le aspettative: che cosa si aspettano in concreto i manager quando decidono di applicare una strategia di CRM?**

Tutti sono concordi con l'identificare l'obiettivo del CRM con l'ottimizzazione del grado di soddisfazione del cliente attraverso la costruzione di rapporti interni all'organizzazione che favoriscano la costruzione di rapporti con la clientela ma gli obiettivi che tracciano le finalità strategiche possono essere molto diversi, ad esempio possono essere obiettivi di vendite o di quote di mercato, possono invece definire una politica aziendale come "una società cliente-centrica" oppure possono definire la volontà di fidelizzare i clienti a rischio di abbandono, ecc.

Le scelte organizzative e tecnologiche sono, di conseguenza, molto diverse tra di loro ma si possono fare due distinzioni generali tra gli obiettivi di gestione efficiente dell'interazione con il cliente e gli obiettivi di *customer intelligence*.

Nella prima classe rientrano tutte le attività che riguardano il contatto diretto con il cliente come la vendita, l'assistenza, le iniziative di marketing, e vengono gestite con l'aiuto di tecnologie di automazione quali la Sales Force Automation, l'automazione delle campagne marketing, ecc. In questa classe di obiettivi viene posta molta enfasi sull'efficienza dei processi, l'azienda deve essere in grado di reagire in modo coerente e qualitativamente alto ad ogni possibile richiesta del cliente. Risulta pertanto molto importante la gestione di tutti i canali di interazione come il punto vendita, il sito web, la posta elettronica, ecc. ed inoltre è necessario che l'azienda si rapporti con l'esterno come un'entità unica e coerente. Tutti i processi aziendali, anche quelli che apparentemente sono più distanti dal cliente, devono essere centrati su di esso, l'azienda deve ad esempio essere in grado di rispondere a qualsiasi e-mail che chieda di nuove tariffe speciali così come deve saper informare tempestivamente sullo stato di avanzamento di un ordine o di un pagamento.

Gli obiettivi che riguardano la *customer intelligence* invece sono rivolti all'interno dell'azienda ma si basano su i dati e le informazioni reperite dalle transazioni effettuate dai clienti con l'azienda e servono per capire veramente chi è il cliente. Le tecnologie impiegate sono di datawarehousing per organizzare i dati per fini di analisi, e tecnologie di modellazione e analisi.

Per capire chi sono i clienti e come si comportano le possibilità di analisi si estendono dall'analisi descrittiva di tipo multidimensionale a quella esplorativa attraverso l'applicazione di tecniche di data mining. Gli obiettivi e i risultati sono diversi, mentre nella prima tipologia l'obiettivo iniziale è conosciuto e il risultato è solitamente una descrizione del fenomeno analizzato (che può essere anche molto articolata), con il *data mining* spesso si parte applicando degli algoritmi sui dati senza avere chiaro che cosa si sta cercando, il risultato di una tale attività è spesso di difficile interpretazione e richiede la stretta interdipendenza tra gli analisti e i manager del marketing.

L'analisi multidimensionale serve per esplorare i dati presenti in azienda (contenuti nel datawarehouse), fotografare lo stato delle cose ed eventualmente approfondire alcuni aspetti ritenuti importanti, risponde a domande come "Chi sono i clienti che spendono di più il sabato pomeriggio nei negozi del nord-ovest?". Questo tipo di analisi è usato per confermare o meno, a livello statistico e quantitativo, le intuizioni e le percezioni dei fenomeni e risponde a domande come "Quanti anni hanno, di quale sesso sono i nostri clienti?".

Infine è una tipologia di analisi che può essere usata per misurare quantitativamente un fatto, ad esempio "Quanto hanno generato in termini di acquisti i clienti a cui ho inviato un regalo".

Per i manager dare una risposta statisticamente valida e significativa a queste tipologie di domande è di grande importanza.

Le analisi di data mining sono condotte invece senza avere deciso a priori su che cosa concentrare l'attenzione, possono servire per raggruppare e classificare elementi, ad esempio i clienti, che hanno delle caratteristiche in comune oppure possono servire per cercare eventi ricorrenti che individuano delle regole, in generale il data mining serve per cercare relazioni non note a priori tra grandi quantità di dati.

Questo tipo di analisi può portare a risultati di difficile lettura e comprensione, nella fase di interpretazione dei risultati del data mining è necessaria la collaborazione tra gli analisti e i manager per validare e dare significato ai risultati ottenuti nel proprio contesto di business.

Negli ultimi anni si è parlato molto di data mining in ambito aziendale e con molto entusiasmo, effettivamente tutto il processo di data mining, dalla preparazione dei dati all'interpretazione dei risultati, apporta una notevole dose di conoscenza sui propri



clienti e sui loro comportamenti passati e futuri ma apporta anche conoscenza sull'azienda stessa perché aiuta i manager a riflettere sui dati da un punto di vista diverso.

Per concludere, le due facce del CRM, CRM operativo e analitico, dovrebbero convivere e supportarsi l'un l'altra, ma nella realtà possono anche essere applicate separatamente.

### **1.3.1 Alcuni esempi eccellenti**

Esiste un grande clamore intorno al CRM, su di esso circolano molteplici definizioni e numerosissime soluzioni informatiche. Per chi è interessato a capire di che cosa si tratti e soprattutto se è ciò di cui ha bisogno la propria azienda è indispensabile sapere che in letteratura e nelle riviste specializzate c'è un certo grado di confusione tra ciò che rappresenta informazione e ciò che è pubblicità di un prodotto o di una consulenza strategica.

Al di là della pubblicità e delle mode e al di là di ciò che viene riportato dalla stampa e dalla letteratura specializzata, i manager che decidono avvicinarsi alla strategia di CRM possono innanzitutto prendere spunto dalle loro esperienze personali come consumatori, esistono infatti esempi eccellenti di applicazioni di CRM i cui risultati sono sotto gli occhi di tutti e a cui i manager si rivolgono e si ispirano; *Tesco, Coop Italia, Zara, Ermenegildo Zegna*<sup>3</sup>.

Tutte le aziende sopra citate rappresentano esempi in cui i sistemi di CRM hanno avuto un grande successo ed hanno portato ottimi ritorni economici e di immagine e nello stesso tempo dimostrano che il CRM può essere applicato in modi molto diversi tra di loro. Coop Italia, ad esempio, sfruttando la sua peculiarità di impresa cooperativa ha introdotto una carta socio magnetica che il cliente (nonché socio della cooperativa) presenta al momento del pagamento della spesa; attraverso un meccanismo di accumulo di punti e di accesso a particolari promozioni la carta socio coop è il maggior vettore del programma di fidelizzazione dell'impresa. Dal punto di vista dell'azienda, la carta socio

---

<sup>3</sup> L'autore cita esempi di approcci diversi al CRM di aziende note di cui è venuto a conoscenza per esperienza diretta come consumatore e di cui i media specializzati, e non, hanno parlato spesso. Non vengono proposti, se non nel caso di Tesco, riferimenti bibliografici precisi e per ulteriori approfondimenti si rimanda ai siti internet di queste aziende e alla ricerca sulla rete internet di argomenti correlati.

serve per associare un profilo anagrafico agli acquisti, in modo tale che si possano fare analisi non solo sugli acquisti in maniera anonima ma anche analisi per conoscere e capire meglio il cliente; dai risultati delle analisi gli esperti prenderanno spunto per ideare nuove attività di marketing e il cliente verrà di nuovo coinvolto nel processo di osservazione-analisi-azione.

Un altro esempio eccellente che ha notevole somiglianze con quello di Coop Italia, è quello di Tesco che è uno delle più grandi catene inglesi di supermercati, Tesco è presente nel settore dalla fine de l'800 e inizialmente era un'impresa cooperativa.

L'esperienza di Tesco nel CRM ha avuto un successo enorme ed è caratterizzata da un programma di fedeltà veicolato attraverso l'uso di una carta *loyalty* chiamata *Clubcard*. Il programma *Clubcard* è stato introdotto all'inizio degli anni '90 quando questo tipo di attività di marketing non era ancora molto diffuso ma anzi era del tutto in via di sperimentazione, la carta *loyalty* ha segnato e trasformato profondamente i comportamenti e le abitudini dei consumatori da un parte e dell'azienda *retail* dall'altra. Un esempio significativo del cambiamento del rapporto tra cliente ed azienda è dato dal fatto che Tesco, prima dell'introduzione della programma *Clubcard*, era uno dei più grandi acquirenti di spazio pubblicitario televisivo, dopo due anni dall'introduzione del programma invece Tesco ha eliminato del tutto la pubblicità in televisione nel periodo natalizio (periodo in cui veniva maggiormente sfruttato il canale) privilegiando il contatto diretto dei propri clienti attraverso canali come la posta, il telefono e il negozio stesso.

Nel caso di Zara, una catena spagnola di abbigliamento che ha avuto una notevole espansione negli ultimi anni stabilendo una presenza fortissima in Europa e in tutto il mondo, lo scopo del CRM è capire che cosa il cliente vuole e quali sono le sue preferenze in base ai dati degli acquisti raccolti in maniera anonima. L'efficienza del processo è il punto focale di questa strategia, attraverso l'analisi dei dati sugli acquisti si cerca di capire quali sono le preferenze e le tendenze della clientela, dai risultati delle analisi parte l'ordine alle fabbriche ed infine il prodotto arriva nei punti vendita. Tale processo che coinvolge i punti vendita, il marketing, la produzione e la distribuzione comprende le fasi di raccolta dei dati, analisi, riordino, produzione e distribuzione dura un tempo brevissimo (circa una settimana) e viene ripetuto ad intervalli di tempo altrettanto brevi. Anche in questo caso il cliente è al centro del processo, attraverso la

continua raccolta di dati e meticolosa analisi l'azienda propone sul mercato solo ciò che il cliente ha *espressamente* richiesto.

Esistono segnali che anche nel settore della moda e dell'alta moda, con caratteristiche diverse da quelle delle grandi catene di abbigliamento come Zara, si stia investendo molto in tecnologie informatiche per monitorare il cliente e i suoi acquisti, per accumulare una conoscenza approfondita e dinamica della propria clientela che serva da guida per le proprie azioni di marketing e non solo; uno dei casi più in vista sulla stampa mondiale è quello dei consistenti investimenti in tecnologie per la *business intelligence* fatti negli ultimi anni da Ermenegildo Zegna.



### 3. Un sistema di Business Intelligence per il CRM

Il concetto alla base del CRM, ovvero la centralità del cliente, non è nuovo non solo per le imprese di servizi ma anche per quelle industriali, che cosa lo rende allora un concetto originale e di così grande attrattiva? Grazie ai progressi della tecnologia e in particolare alla diffusione di internet e all'aumento della capacità di calcolo parallelamente alla diminuzione dei costi, il consumatore ha possibilità illimitata di informazione e non ha né vincoli di spazio né di tempo, in definitiva ha acquistato un notevole potere rispetto all'azienda. Le aziende quindi non possono far altro che mettere al centro dei loro processi il cliente per sviluppare servizi e prodotti migliori e personalizzati. Il CRM rappresenta per la maggior parte delle attuali imprese una grande opportunità per rimanere sul mercato.

Gli stessi progressi tecnologici che hanno reso il cliente più forte possono però anche essere sfruttati per aiutare le aziende a rispondere alle nuove esigenze dei clienti e alla crescente competizione ed infatti negli ultimi anni sono nate moltissime soluzioni commerciali per l'implementazione dei sistemi di CRM.

Come già anticipato nel capitolo precedente il CRM si compone di un aspetto operativo che comprende le tecnologie a supporto dell'automazione della forza vendita (SFA), l'automazione del marketing, l'assistenza al cliente, l'e-Commerce e i Call Center e di uno analitico o di *Business Intelligence*. Nella figura 1 viene riportato per scopi di sintesi e di esempio quali sono le esigenze aziendali a cui il CRM risponde.

Mentre le applicazioni di CRM operativo sono rivolte all'interazione diretta con il cliente e servono per tenersi in contatto con la clientela e registrare i dati relativi a tali contatti, i sistemi di CRM analitico sono rivolti all'interno dell'azienda per l'analisi dei dati a supporto delle decisioni del marketing e non solo. Il confine tra queste due tipologie di applicazioni è tuttavia molto sottile e confusa e infatti un sistema di CRM completo comprendo applicativi che coprono entrambi gli aspetti.

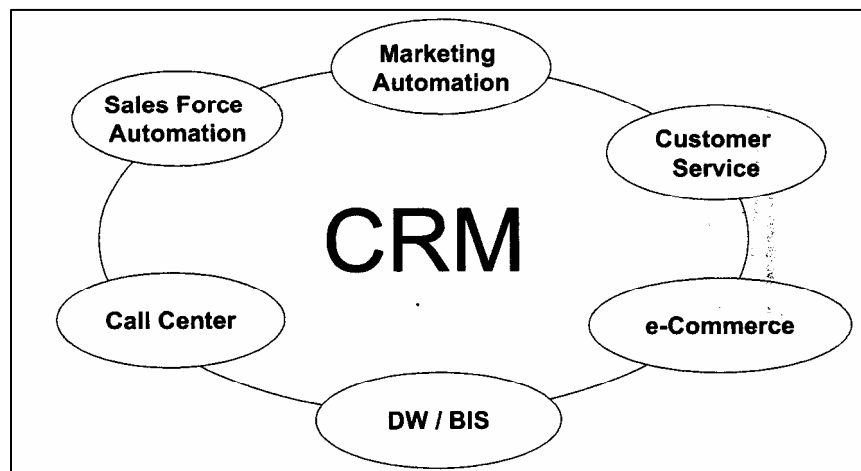


Figura 3: CRM ed esigenze aziendali

Sono infatti molto diffusi anche strumenti che seguono e supportano tutto il processo dalla gestione dell'automazione del canale, alla costruzione del datawarehouse, l'analisi OLAP e di data mining in modo integrato. Questo tipo di soluzioni possono anche essere commercializzate in pacchetti modulari e poi integrabili successivamente.

La soluzione di gestione informatizzata del processo in maniera integrata proposta da questo tipo di strumenti seguirebbe in maniera adeguata la filosofia della strategia di CRM ma presenta degli svantaggi. I costi di acquisto di tali pacchetti è altissimo, i tempi per la parametrizzazione e messa in funzione dello strumento sono molto lunghi e richiedono oltre che personale specializzato nel software in questione anche personale aziendale che conosce le caratteristiche tecnologiche dei sistemi già in uso in azienda. I costi e i tempi sono tali che solitamente un'impresa di piccole o medio-piccole dimensioni non può sostenerli. Una soluzione di questo genere nella realtà si dimostra molto rigida, invece che modellarsi sui processi già esistenti, presenta dei vincoli sulla tipologia dei processi che deve supportare.

In definitiva, soluzioni integrate di *business intelligence* rappresentano investimenti molto onerosi per tempi e costi e devono essere scelti solo dopo aver chiarito i requisiti necessari per l'adozione di una tale soluzione e i benefici che se ne ricaverebbero.

### 3.1 Tecnologie a supporto del CRM operativo

Il CRM operativo comprende tutte le tecnologie che servono per automatizzare i processi operativi al fine di ottimizzare il servizio fornito al cliente, i processi operativi coinvolgono principalmente, ma non solo, l'area del marketing, delle vendite e del servizio al cliente, aree che tradizionalmente non sono perfettamente integrate e comunicanti tra di loro.

La gestione di tutti i canali di contatto, come i punti vendita, gli intermediari di vendita, i *call center*, la posta elettronica, ecc., deve essere fatta con l'obiettivo di memorizzare tutte le interazioni con il cliente mantenendo l'informazione su di esso in maniera tale che la sua storia in azienda sia ricostruibile ed accessibile da tutte le aree e per scopi diversi.

Per una gestione coerente ed omogenea dei canali è necessario disporre di un database dei clienti da cui sia possibile agganciarsi agli altri database aziendali: DB vendite, DB ordini, DB prodotto e altri. Ciò che distingue le soluzioni di CRM operativo dai tradizionali sistemi *ERP* è la centralità del cliente nella gestione dei processi; gli applicativi più comunemente usati come quello per l'assistenza al cliente o l'automazione del marketing girano sul database del cliente e oltre alla finalità di gestione efficiente del processo offrono strumenti e funzionalità per l'analisi e il rapido accesso alle informazioni sul cliente, mentre invece i dati contenuti e trattati in un sistema ERP sono a livello di transazioni, strutturati in unità operative come numeri d'ordine, stock di magazzino, ecc.

Tra le soluzioni *front-office*, merita una descrizione dettagliata l'automazione del marketing perché le soluzioni tecnologiche proposte sono a cavallo tra il CRM operativo e quello analitico.

Il marketing comprende una serie di attività che coinvolgono direttamente il cliente, ad esempio le campagne marketing, le ricerche di mercato, le indagini dirette, queste attività vengono veicolate attraverso i canali di output già citati. Le decisioni di quali attività intraprendere e con quali modalità vengono però prese internamente dai manager del marketing.

Il processo decisionale interno del marketing può avvalersi del sistema di automazione esterno del marketing sia nelle fasi iniziali di *Pre-analisi* e *Progettazione e*

*Pianificazione* dove applicazioni di *Business Intelligence* elaborano i dati raccolti nelle fasi di contatto con il cliente per le analisi di segmentazione e profilatura sia nella fase finale di *Analisi dei Risultati* dove si cerca di misurare l'efficacia della campagna con l'analisi dei dati raccolti sul comportamento del cliente.



**Figura 4:Le fasi del processo di Marketing Automation**

### 3.2 Tecnologie a supporto del CRM analitico

Il processo di CRM analitico, supportato da strumenti di *Business Intelligence*, si compone delle fasi di raccolta delle informazioni sul cliente, analisi delle informazioni, formulazione di una strategia basata sui risultati emersi dall'analisi e infine intervento in base alla strategia.

Le tecnologie che compongono i sistemi di *Business Intelligence* sono tipicamente di datawarehousing, strumenti di analisi multidimensionale, di *data mining* e di *reporting*. Come già detto, solitamente i canali di input dei dati e i canali di output attraverso cui si mettono in atto gli interventi in base alla strategia sono spesso gestiti con applicativi di CRM operativo.



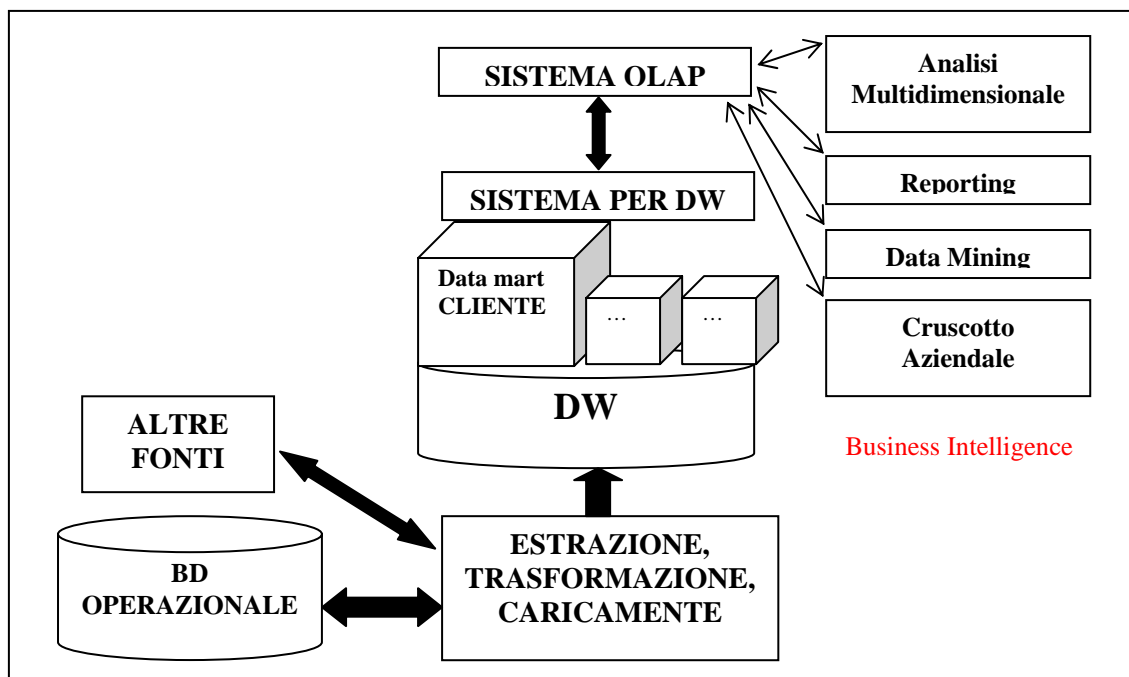


Figura 5: esempio di architettura a supporto per un ambiente di business intelligence

### 3.2.1 Il Datawarehouse

La base di dati su cui poggiano le applicazioni di CRM analitico, che d'ora innanzi chiamerò *datawarehouse* del cliente è una raccolta di dati storici integrati, non volatile, organizzata per temi e finalizzata al recupero di informazioni di supporto ai processi decisionali.

In genere il *datawarehouse* aziendale comprende e collega le diverse aree informative dell'organizzazione, nel caso del CRM la parte di *datawarehouse*, chiamata anche *data mart*, interessante è quella che riguarda il cliente, ma possono esistere molti altri *data mart*, a seconda del tipo di business che caratterizza l'azienda (ordini, magazzino).

Tutti i dati sul cliente esistenti in azienda sono raccolti in un'unica base di dati, le fonti possono essere sia interne sia esterne, alcuni esempi sono: informazioni di tipo qualitativo su alcuni clienti o classi di clienti (ad esempio clienti VIP), dati provenienti da altre basi di dati operazionali come quella delle vendite o dei prodotti, dati (qualitativi o quantitativi) provenienti da ricerche di mercato e indagini socio-demografiche eseguite dall'azienda stessa o comprate da fonti terze. Vista l'eterogeneità delle fonti sia da un punto di vista di contenuto che di forma, il lavoro necessario per

creare un'unica base di dati che organizza tutti i dati sul cliente in maniera coerente ed integrata è molto lungo e complesso.

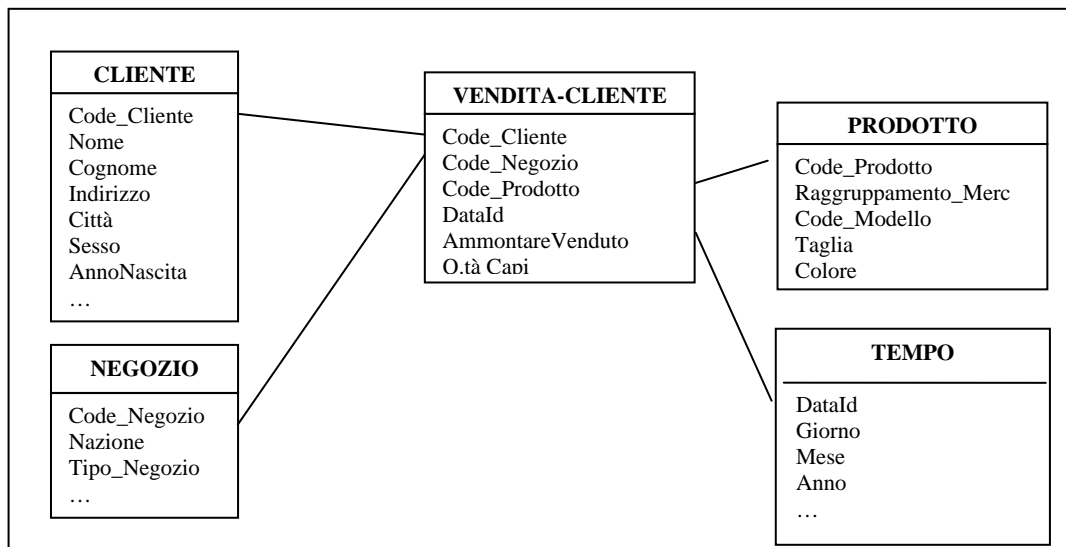
Oltre ad integrare le diverse fonti di dati sul cliente il datawarehouse deve mantenere le informazioni nel tempo, deve cioè permettere di ricostruire la storia delle interazioni tra il cliente e l'azienda, la dimensione temporale è cioè di fondamentale importanza.

Il *datawarehouse* deve, in sintesi, essere organizzato al fine di rendere possibile analizzare i dati ed elaborare informazioni sul cliente e sul suo comportamento; deve cioè essere accessibile e facilmente interrogabile negli aspetti ritenuti più importanti. La logica con cui si guardano i dati è molto diversa da quella dei database transazionali, nel *datawarehouse* si osserva un fatto e lo si misura prendendo in considerazione punti di vista diversi e più di un punto di vista contemporaneamente (analisi multidimensionale).

Il *datawarehouse* è progettato e sviluppato da un team congiunto di esperti di *database*, analisti e manager del marketing proprio a causa del carattere di estrema funzionalità rispetto alle esigenze di analisi che deve avere il *datawarehouse* nel concetto di CRM.

Le esigenze di analisi degli utenti del datawarehouse si concretizzano in interrogazioni al *database* non definite a priori, caratterizzate da un ridotto insieme di transazioni che operano sulla base di dati, ciascuna delle quali però richiede l'accesso ad un elevato numero di righe ed implica una complessità delle *join* che potrebbe avere un impatto molto negativo sulle performance del sistema. La logica dell'organizzazione delle basi di dati operazionali, tipicamente relazionali, che si basa su operazioni predefinite relativamente semplici e che coinvolgono pochi dati di dettaglio, non può pertanto essere adatta a quella del datawarehouse.

Il modello multidimensionale viene implementato nei sistemi relazionali da uno schema asimmetrico, chiamato schema a stella, dove c'è una tabella dominante di grandi dimensioni (tabella dei fatti, in figura 6 rappresentata da VENDITA-CLIENTE) che è l'unica ad avere collegamenti multipli (chiavi esterne) con le altre tabelle (tabelle delle dimensioni, CLIENTE, NEGOZIO, PRODOTTO, TEMPO), le tabelle delle dimensioni sono in *join* solo con la tabella dei fatti.



**Figura 6: un esempio di schema a stella**

Questo tipo di organizzazione riflette la visione che i manager e l'analista hanno del business e pertanto minimizza il numero di *join* necessarie per ciascuna interrogazione. Inoltre per motivi di efficienza si può decidere di memorizzare dei dati calcolati a partire dai dati di dettaglio, escludendo o meno il dettaglio dal datawarehouse, la scelta di quali dati di sintesi precalcolare e memorizzare deve essere fatta in base alle specifiche dell'analista che esprimerà le sue priorità in base alla frequenza di accesso alle informazioni.

Esistono anche dei sistemi specializzati, chiamati MOLAP (*Multidimensional OLAP*), che implementano direttamente il modello multidimensionale ma sono adatti per la loro semplicità a situazioni in cui le dimensioni rilevanti per l'analisi di un certo fenomeno sono contenute, quando le dimensioni crescono, come spesso succede quando con l'esperienza si capiscono le potenzialità del datawarehouse, la complessità del modello diventa ingestibile e solitamente si passa ad un'organizzazione ibrida di diversi modelli. La creazione di un datawarehouse è un processo molto lungo e costoso (almeno 1 anno), ed è molto delicato soprattutto perché si devono incontrare motivi di efficienza tecnica (minimizzazione di memoria utilizzata, limiti e vincoli di capacità di calcolo e di reperibilità dell'informazione) con le esigenze degli utenti di flessibilità, accessibilità e possibilità di aggiornamento in tempo reale. Inoltre, poiché nel tempo le esigenze manageriali e le caratteristiche del business possono cambiare, la progettazione e lo

sviluppo del datawarehouse assumono un carattere incrementale e diventa difficile prevedere la loro durata.

Nel capitolo 4 un caso pratico illustrerà le problematiche della progettazione del datawarehouse e una possibile soluzione.

Il datawarehouse è organizzato per fini analitici e gli scopi con cui si conducono le analisi possono essere diversi; le possibilità di analisi si estendono dall'analisi descrittiva di tipo multidimensionale a quella esplorativa attraverso l'applicazione di tecniche di data mining.

### **3.2.2 L'Analisi Multidimensionale**

Le informazioni che scaturiscono dall'elaborazione dei dati del datawarehouse attraverso analisi di tipo multidimensionale sono destinate ai manager e servono per supportarli nelle loro decisioni.

Le informazioni *direzionali* sono organizzate per temi o fatti e forniscono degli indicatori che servono per misurare le performance aziendali secondo prospettive rilevanti (dimensioni).

Gli indicatori servono quindi per misurare in termini quantitativi un fatto, senza però perdere di vista la globalità del business e dell'azienda. Il livello di sintesi dell'informazione da presentare ai manager è solitamente molto alto, le informazioni che servono loro si limitano a pochi indicatori che descrivono un fenomeno e l'impatto che questo ha sull'intera struttura aziendale; la sintesi però deve essere in grado di mettere in luce le criticità e una volta che queste sono state individuate deve essere possibile scendere nel dettaglio e approfondire.

Le elaborazioni analitiche sul datawarehouse sono fatte con l'uso della tecnologia OLAP (*On-Line Analytical Processing*), che permette di generare riassunti descrittivi o comparativi dei dati. I dati provenienti dal datawarehouse sono pertanto utilizzati per descrivere, in modo esauriente ed approfondito, la realtà del business nei suoi vari aspetti.

Le elaborazioni di tipo OLAP devono permettere di avere una visione multidimensionale dei dati e di navigare agevolmente tra i diversi livelli di sintesi.

La visione multidimensionale e dinamica dei dati può essere rappresentata in una forma tipo cubo (nel caso in cui si prenda in considerazione tre dimensioni) invece che in forma tabellare. In figura 4 è mostrata la rappresentazione di un cubo a partire dallo schema a stella del *datawarehouse* di figura 6, dove la misura è *Ammontare Venduto*, il cubo ha come dimensioni il tempo indicato da T, le tipologie di prodotto, P, e i negozi M. Il fatto osservato è la vendita-cliente, ogni casella bianca indica l'ammontare venduto in un certo periodo di tempo, in un certo negozio, di un determinato prodotto. La figura 7 rappresenta un cubo esteso, dove cioè vengono calcolate ed eventualmente memorizzate delle misure di sintesi. Le caselle grigio chiaro rappresentano la somma dell'ammontare venduto su una dimensione tenute costanti le altre, come evidenziato in figura 4, *totale vendite di P4 in M1* indica la somma dell'ammontare delle vendite in ciascun periodo del prodotto P4 nel negozio M1. Le caselle grigie rappresentano un ulteriore livello di sintesi, nella figura 7 *totale vendite nel negozio M1* indica la somma dell'ammontare delle vendite di ciascun periodo di tutti i prodotti nel negozio M1 ed infine la casella grigio scuro rappresenta il valore venduto di tutti i prodotti in tutti i negozi in tutto l'intervallo di tempo preso in considerazione.

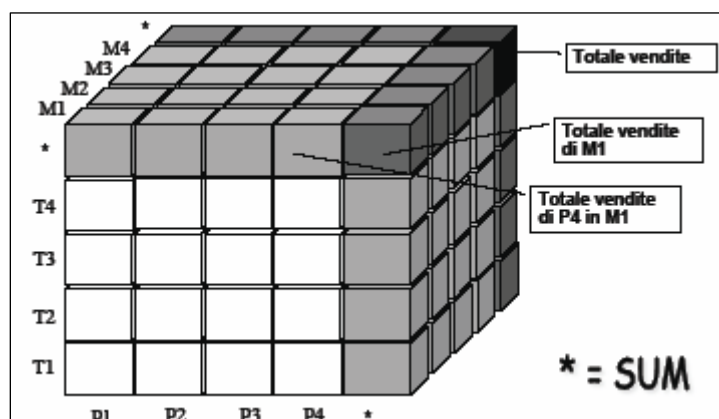


Figura 7: il cubo esteso

Un'altra particolarità del modello dimensionale è l'uso delle gerarchie, una dimensione può essere descritta da un insieme di attributi che hanno tra di loro una relazione gerarchica, ad esempio la relazione geografica stato-regione-comune; navigare la gerarchia, salendo e scendendo nella relazione equivale ad aggregare o disaggregare i dati.

Questa tipologia di analisi può essere realizzata interrogando direttamente il datawarehouse, il tipo di interrogazioni dipende da come sono stati organizzati i dati del datawarehouse, se si sono mantenute le logiche di organizzazione relazionale o si è passati alla logica multidimensionale.

Nel caso di operazioni OLAP su dati multidimensionali l'SQL (*Standard Query Language*) è stato esteso con funzioni che rendono più efficienti e agevoli le operazioni tipiche delle elaborazioni OLAP, che, come si è visto nell'esempio del cubo esteso, sono operazioni di aggregazione della misura della tabella dei fatti su una o più dimensioni tramite funzioni come ROLL-UP e CUBE.

L'analisi multidimensionale può essere anche molto sofisticata e prendere in considerazione misure di tipo statistico sui dati, alcuni DBMS come Oracle e Red Brick estendono l'SQL per l'analisi dei dati con delle funzioni analitiche. Le funzioni analitiche sono di tipo statistico (media, deviazione standard, *ranking*) applicate a dei contesti particolari come la media mobile calcolata in un intervallo di tempo o come trovare il rango (*rank*) di una misura in un gruppo.

Questo tipo di analisi è molto comune e il ricorso alle funzioni analitiche risolve sia problemi di efficienza sia problemi di complessità nella scrittura delle interrogazioni SQL.

Solitamente l'analista non interroga direttamente i dati ma è dotato di un'interfaccia che offre sia la possibilità di accesso ai dati sia di analisi, gli strumenti commerciali più usati nelle aziende offrono in un'unica soluzione funzionalità di accesso ai dati, di elaborazione ed analisi e di *reporting*. Nel capitolo 4 saranno descritte nel dettaglio le caratteristiche di uno di questi strumenti, *BusinessObjects*.

Un sistema per analisi OLAP fornisce quindi gli strumenti per generare rapporti descrittivi e di sintesi a partire da interrogazioni mirate e specifiche; vista la complessità e la numerosità dei dati presenti in azienda il successo di una tale operazione non è affatto scontato.

L'analisi multidimensionale deve dare una panoramica approfondita della realtà, dei problemi e delle opportunità del business ma deve anche essere sintetica ed efficace. Oltre alle problematiche tecniche e tecnologiche la fase dell'analisi coinvolge anche problematiche legate alla conoscenza del contesto organizzativo e del settore in cui l'azienda opera.

L'analista congiuntamente con il team di progetto del CRM deve, saper trasformare i dati in informazioni e le informazioni in conoscenza utile ai manager per prendere le decisioni, come vedremo nei capitoli successivi, l'individuazione degli indicatori più appropriati ed *espressivi* è un processo che oltre a coinvolgere persone con competenze e ruoli diversi è ciclico ed incrementale.

Se i risultati delle analisi sono ben preparati e mirati per rispondere alle esigenze dei manager, i risultati susciteranno nuovi interessi e curiosità e dalle criticità evidenziate scaturiranno approfondimenti.

L'ultimo passo dell'analisi è la presentazione dei risultati, l'efficacia dell'intero processo dipende infine dalla capacità di saper comunicare e condividere la conoscenza accumulata.

Gli strumenti commerciali di analisi offrono funzionalità di reporting anche molto sofisticati, solitamente si sceglie uno standard di presentazione che viene riapplicato alle presentazioni successive in maniera tale che il documento di *reporting* oltre che ad essere percepito come familiare e distintivo sia anche confrontabile con quelli precedenti e con quelli successivi.

### **3.2.3 Il Data Mining**

E' possibile estendere le capacità esplorative e conoscitive dell'analisi dei dati attraverso l'applicazione di tecniche di *data mining* partendo da un punto di vista diverso rispetto a quello dell'analisi multidimensionale; il *data mining* consiste infatti nella ricerca in una base di dati di associazioni e relazioni non note a priori.

Il data mining è un processo e consiste nelle fasi di selezione, esplorazione e modellazione di grandi masse di dati e poiché ha lo scopo di scoprire regole non note per ottenere un vantaggio di business comprende anche le fasi di interpretazione e valutazione rispetto al contesto del business.

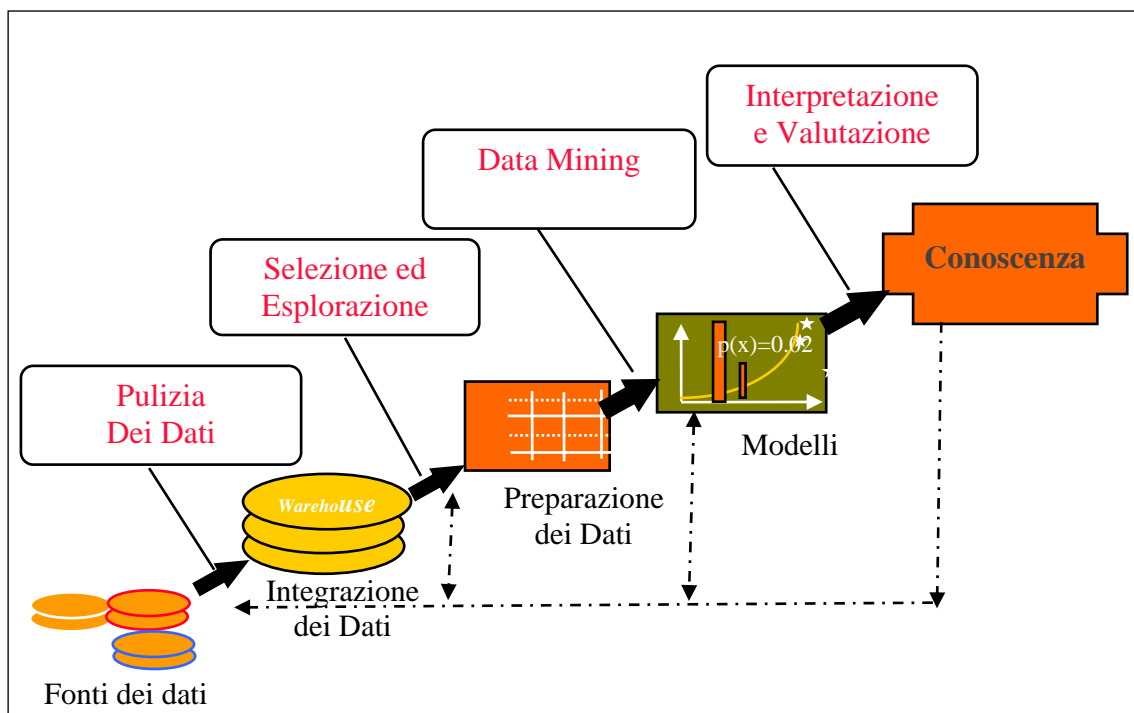


Figura 8: il processo di KDD

A monte del processo di KDD (*Knowledge Discovery in Data bases*) il primo passo è di definire gli obiettivi con cui si intraprende il processo e trasformare il problema di business emerso dagli obiettivi in un problema di analisi quantitativa.

Gli obiettivi devono essere definiti in modo tale che siano chiari, specifici e misurabili, ad esempio *identificare i clienti che sono a rischio di abbandono e descrivere le caratteristiche che rendono simili questi clienti*.

Le varie fasi del processo di KDD sono caratterizzate da una serie di operazioni descritte di seguito.

1. Pulizia dei dati: le fonti da cui attingere i dati interessanti per l'analisi sono molteplici, nel nostro caso supponiamo che esista in azienda un datawarehouse che contiene ed integra tutte le fonti di dati.
2. Selezione dei dati appropriati per l'analisi: in questa fase si deve decidere quali e quanti dati tra quelli disponibili trattare nell'analisi. Questa operazione è molto delicata perché dipende fortemente dallo stato e dalla quantità di dati a disposizione e perché non esistono regole precise che guidano l'analista nel fare la scelta migliore. Per quanto riguarda la numerosità l'insieme dei dati deve



contenere i dati necessari per descrivere i vari aspetti del fenomeno, anche quelli più marginali. Inizialmente è bene non escludere a priori nessuna variabile sulla base della conoscenza intuitiva e aleatoria del business ma aspettare che siano i dati stessi a rivelare ciò che è importante o meno. Riguardo alla storicità, nel caso in cui esista un aspetto di stagionalità nel business, i dati devono riguardare più di una stagione, in generale poiché il processo si basa sull'apprendimento dalle esperienze passate è necessario avere i dati sui comportamenti passati dei clienti e dell'azienda.

3. Esplorazione: prima di iniziare a costruire il modello è indispensabile conoscere i propri dati e il contesto di business. Le analisi di tipo multidimensionale sono di grande aiuto per osservare i dati da diversi punti di vista e offrono delle sintesi interessanti per approfondire la conoscenza del business. E' inoltre molto utile che per ogni variabile vengano calcolate la frequenza dei valori, la media, la mediana e vengano eseguite tutta una serie di operazioni di controllo e visualizzazione dei dati per assicurarsi che i dati siano trattabili.
4. Creare un modello: un modello, nella sua accezione più generale, è una rappresentazione concettuale che descrive e spiega le regole con cui la realtà funziona che può essere usato per fare delle inferenze su come andranno le cose nel mondo reale.

In questo contesto viene costruito un modello sui dati decidendo in base agli obiettivi di analisi prefissati quale strategia di DM mettere in atto. Esistono tre strategie, quelle supervisionate, quelle non supervisionate e le analisi delle affinità; le strategie supervisionate hanno l'obiettivo di costruire un modello di rappresentazione con un apprendimento controllato da casi in cui è nota la classificazione (valore dell'attributo *target*). Le strategie non supervisionate invece rappresentano i dati con un apprendimento non controllato da casi in cui non è nota la classificazione, tutti gli attributi sono indipendenti ed a priori non sono note di quante e di che tipo di classi sarà l'output. Infine, le analisi delle affinità applicano una strategia in cui si cercano relazioni interessanti tra un insieme di valori.

Ogni tipologia di strategia è implementata da una serie di tecniche diverse, a loro volta implementate da algoritmi diversi, ad esempio la strategia supervisionata

può essere implementata usando la tecnica della Classificazione (implementata a sua volta da algoritmi come il CHAID o CART); un esempio di tecniche per una strategia non supervisionata è la *Cluster Analysis* (implementata dal *Simple K-means*).

L'insieme dei dati che costituisce il modello è diviso in tre sottoinsiemi (solitamente disgiunti): il *training set* è usato per costruire un insieme di modelli, il *validation set* è usato per scegliere il miglior modello da implementare e infine il *test set* serve per determinare le performance del modello con dati nuovi.

5. Scelta del modello finale: in questa fase oltre a quanto detto sopra nella fase di costruzione del modello concorrono alla scelta finale considerazioni sugli obiettivi prefissati e i vincoli di business sia in termini di budget che di tempo. La scelta di quale tecnica o combinazioni di tecniche utilizzare dipende dal tipo di problema e dal tipo di dati disponibili, nel capitolo 5 verrà descritto il legame tra un ambito applicativo e la scelta della migliore tecnica corrispondente.
6. Valutazione del modello: nella valutazione di un modello si prendono in considerazione le caratteristiche tecniche di un modello applicato come l'accuratezza, il livello di sicurezza sull'attendibilità dei risultati, ma anche considerazioni come la capacità espressiva, la semplicità e la comprensibilità del modello. I parametri usati per misurare l'accuratezza di un modello o il suo grado di confidenza variano a seconda del tipo di modello implementato, così come varia l'importanza attribuita alla semplicità e comprensibilità del modello e dei suoi risultati a seconda degli obiettivi iniziali.
7. Interpretazione dei risultati: interpretare i risultati dell'analisi del *data mining* è un momento critico del processo di analisi, sia perché i risultati sono di difficile lettura e comprensione sia perché l'output del *data mining* non dà informazioni sempre utili o significative per il business. In questa fase l'analista deve condividere e illustrare i risultati al manager, attraverso la collaborazione tra queste due figure il processo culmina con il consolidamento della conoscenza acquisita durante tutte le fasi del processo.
8. Implementazione delle regole nei processi decisionali: l'integrazione dei risultati nei processi decisionali aziendali è il passo finale del processo e coinvolge i manager.

In generale il processo di analisi con il *data mining* apporta una grande quantità di informazioni, dalla fase iniziale di preparazione dei dati a quella finale di interpretazione, anche nel caso in cui non emergano aspetti particolarmente sorprendenti, aiuta i manager e l'analista ad osservare e ragionare sui dati da un punto di vista diverso.



## 4. Sviluppo di un progetto di CRM

### 4.1 Contesto di business

L'Azienda in questione è un'azienda italiana che opera nel campo della moda, è stata fondata nel 1975 ed è tuttora guidata dallo stesso imprenditore che ne è sia proprietario che amministratore.

Fin dai primi anni l'Azienda ha stabilito una presenza in altri paesi, inizialmente in Europa e Stati Uniti per poi espandersi in tutti i continenti. La struttura organizzativa attuale ha una forma divisionale in cui le divisioni estere godono di una certa autonomia amministrativa e strategica rispetto a quella principale italiana, la presenza dell'imprenditore determina un'unità e coerenza di fondo dell'immagine dell'azienda nel mondo.

L'Azienda è nata come impresa esclusivamente commerciale devolvendo all'esterno la produzione e la distribuzione ma con il tempo ha integrato tutte le componenti della *supply chain*, acquisendo o acquistando quote partecipative delle industrie produttive, collocate quasi prevalentemente nel territorio italiano, delle società di distribuzione sparse in tutti i continenti e stringendo delle *joint ventures* con società e personalità di particolare prestigio e di fama internazionale.

Le attività commerciali si dividono tra la vendita *wholesale* e quella nei negozi monomarca; una parte dei negozi monomarca è di proprietà del gruppo.

Questa azienda è stata la prima nel panorama della moda e dell'alta moda a diversificare l'offerta in modo da conquistare mercati diversi, la diversificazione è avvenuta sul prodotto, non solo abbigliamento e accessori ma anche arredamento, ristoranti, bar e alberghi per promuovere un *life style* pervaso dalla marca. La diversificazione è avvenuta anche su segmenti di clientela diversi: la linea principale di abbigliamento è la prima disegnata dall'imprenditore ed è quella che detta e supporta lo stile e l'immagine dell'intera azienda, è destinata ad un pubblico con grandi capacità di acquisto e di età media, successivamente viene lanciata una seconda linea destinata prevalentemente alla vendita nei negozi multimarca, che ha un prezzo inferiore ma mantiene lo stile della prima. Dopo pochissimi anni si aggiunge una linea per un pubblico più giovane con

prezzi più contenuti ma di ottima qualità, poi una linea per bambini, una per giovanissimi e una dedicata solo ad alcuni paesi.

La diversificazione spinta unita però a una forte identità dell'immagine proposta è stata la chiave del successo dell'azienda che negli ultimi venti anni ha avuto un'espansione fortissima.

L'azienda si è distinta nel suo settore per avere delle forti competenze e capacità interne che hanno creato e offerto sul mercato un prodotto di altissima qualità e grandissimo fascino.

L'Azienda si è distinta anche per le grandi capacità manageriali dell'imprenditore e del suo staff che hanno saputo dare una direzione economica e finanziaria stabile.

Ma al di là delle grande creatività, l'Azienda ha spiccato sulle altre per aver fatto le scelte competitive migliori in maniera tempestiva, innovativa ed intelligente.

L'Azienda può essere definita un'impresa proattiva e lo è stata fin dai suoi inizi, proattiva ed orientata alla competizione. Le scelte e le mosse competitive sono state supportate dalla capacità di produrre un grande prodotto ma soprattutto dalla conoscenza del settore e del mercato e delle regole che li governano.

La storia delle sue manovre competitive parte con un'idea creativa, per realizzare tale idea invece che di occuparsi dell'intero ciclo di produzione l'imprenditore ha dato in *outsourcing* tutte le fasi della produzione e della distribuzione e si è concentrato su ciò che sapeva fare e sui mezzi che aveva a disposizione.

Il prodotto si distingue per qualità e bellezza ma soprattutto perché possiede uno stile definito e riconoscibile, è pervaso da un'idea di fondo, comunica quasi una filosofia di vita. Una volta affermatosi il prodotto e il suo stile, lo stile si è declinato in una serie sempre più ampia di proposte differenti senza mai perdere la propria essenza ed immagine di fondo; queste mosse strategiche hanno portato alla definizione di una nuova tendenza, il *life-style* dettato dal marchio.

Ed ora? La strada della diversificazione è stata battuta con successo e la nuova evoluzione è quella di guardare al cliente e non solo ai competitori e alle regole tradizionali del mercato.

La diversificazione spinta, non a caso, sfocia nella personalizzazione e, all'estremo, nel marketing *one-to-one*, la nuova chiave per il successo competitivo è conoscere il nuovo cliente.

Vista la complessità di entità che interagiscono con l'Azienda e vista la complessità del business lo sfruttamento delle capacità offerte dalla tecnologia di comunicazione ed integrazione era inevitabile.

Tra i progetti più importanti c'è l'integrazione della *supply chain*, nelle fasi a monte della vendita al consumatore finale: nella fase di interazione con le industrie produttive e nella fase di vendita ai negozi monomarca e multimarca.

Il progetto di CRM e i sistemi tecnologici ed informatici coinvolti nel progetto servono per arrivare al cliente finale e per chiudere la catena.

L'Azienda ha un particolare rapporto con la tecnologia, molto dinamico ed aperto ad ogni opportunità che l'innovazione tecnologica può portare. La complessità del sistema informatico ha portato l'Azienda a cercare le competenze necessarie fuori praticando la formula dell'*outsourcing* in maniera mitigata, i sistemi e le persone lavorano dentro l'Azienda e solo le competenze, se necessario vengono prese da fuori (consulenti). Questa scelta permette un controllo forte sullo stato dei sistemi e sulle scelte progettuali ed implementative. In Azienda si è capita l'importanza della tecnologia informatica e l'impatto che questa può avere sull'organizzazione, dei cambiamenti che può provocare a partire da quelli apportati all'organizzazione della mansione ai cambiamenti che porta nella gestione di interi processi; pertanto i sistemi informativi si occupano esplicitamente anche di organizzazione.

Nei contesti dove si cerca di far collaborare armoniosamente persone con competenze diverse sono indispensabili figure di collegamento che sappiano avvicinare e far comunicare due mondi diversi ed in particolare in questo capitolo si parlerà del ruolo dell'analista dei dati per il CRM.

## **4.2 Il progetto di CRM**

L'immagine dell'azienda presente, anche se con sfumature diverse, in tutti i suoi diversi marchi è riconoscibile in tutto il mondo, il carisma di questa immagine fa sì che spesso l'acquisto da parte del cliente sia principalmente guidato da emozioni e desideri di identificazione con lo stile del marchio e solo secondariamente da motivi strettamente

legati alla qualità e alla bellezza del prodotto. Per molti anni è stata l'azienda che imponeva al mercato il prodotto e il consumatore che vi si doveva adattare, negli ultimi anni la situazione è cambiata soprattutto perché sono cambiati i consumatori. Per adattarsi al cambiamento si deve innanzitutto capire e conoscere i propri clienti.

L'azienda ha sempre avuto un canale diretto con il consumatore finale attraverso i negozi di proprietà ed ha sempre portato avanti una strategia di attenzione e cura verso il cliente. Il cliente per l'azienda però era riconoscibile solo in parte, la grande massa dei clienti era anonima e solo ad una piccola parte di essi, conosciuta e riconoscibile, era destinato un trattamento speciale.

Le attività di marketing si svolgevano essenzialmente su due piani, uno di pubblicità attraverso i canali dei media, la pubblicità è usata per promuovere un prodotto o una nuova collezione ma soprattutto è sempre usata per diffondere e rafforzare l'idea e lo stile del marchio. L'altro piano è il contatto diretto con attività di *public relation* sfruttando essenzialmente il canale del negozio, le public relation si basano sulla conoscenza diretta e personale del cliente e per la loro natura sono destinate ad un gruppo ristretto e selezionato di persone, hanno lo scopo di costruire una relazione duratura e profittevole con questi clienti ma anche in questo caso l'impatto più forte è sull'immagine dell'azienda; i clienti vengono selezionati perché in grado con i propri comportamenti di influenzare il loro ambiente di riferimento, si tratta quindi di personaggi famosi o VIP ma anche di persone che a livello locale godono di una certa visibilità e prestigio.

L'azienda, pur operando in un settore di nicchia e vivendo un momento molto positivo, ha risentito dei cambiamenti avvenuti a livello globale e si sta attrezzando per adeguarvisi.

Con il crescere delle dimensioni e della complessità l'azienda ha posto una certa enfasi sull'informatizzazione e l'automatizzazione dei processi e conseguentemente sull'importanza del ruolo che i dati e le informazioni quantitative devono avere nei processi decisionali.

L'azienda ha ancora grandi capacità di attrattiva e di seduzione nei confronti dei clienti ma ha comunque deciso di seguire la tendenza generale di aprirsi al cliente per cercare di capire chi è e che cosa vuole; per fare questo è agevolata dal fatto di disporre del contatto diretto principale e tradizionale, il negozio.



Il progetto di un sistema informatizzato del CRM è relativamente recente, la fase del progetto pilota ha coinvolto solo una parte dei negozi in alcuni paesi europei, in particolar modo i negozi italiani di proprietà del gruppo per ragioni di facilità di azione e di controllo.

Il progetto coinvolge principalmente l'area del marketing, l'area *retail* e i sistemi informativi, anche nella sua fase pilota è un progetto di lunga durata che richiede un certo impegno finanziario e di risorse umane con ruoli, capacità e competenze diverse.

Il progetto è partito in forma di pilota perché inizialmente non era affatto scontato che l'azienda avesse bisogno di un approccio tipo CRM e che esperienze di CRM di successo in settori molto diversi potessero essere riproposti nel settore della moda. Il progetto pilota doveva quindi dimostrare l'applicabilità e la validità della strategia di CRM nell'azienda e la coerenza rispetto alle regole del business.

Gli obiettivi del progetto pilota erano quelli di dare un nome ai clienti costruendo un'infrastruttura tecnologica per reperire e memorizzare i dati sul cliente e sul suo comportamento d'acquisto e di costruire un ambiente di *datawarehousing* che permetta di fare delle analisi sui dati raccolti. La fase di reperimento dei dati è lunga e molto costosa e di per sé non porta risultati significativi che giustificano l'investimento, i risultati delle analisi su tali dati dovevano servire per confermare l'utilità del progetto intrapreso nel suo complesso e per giustificare la continuazione.

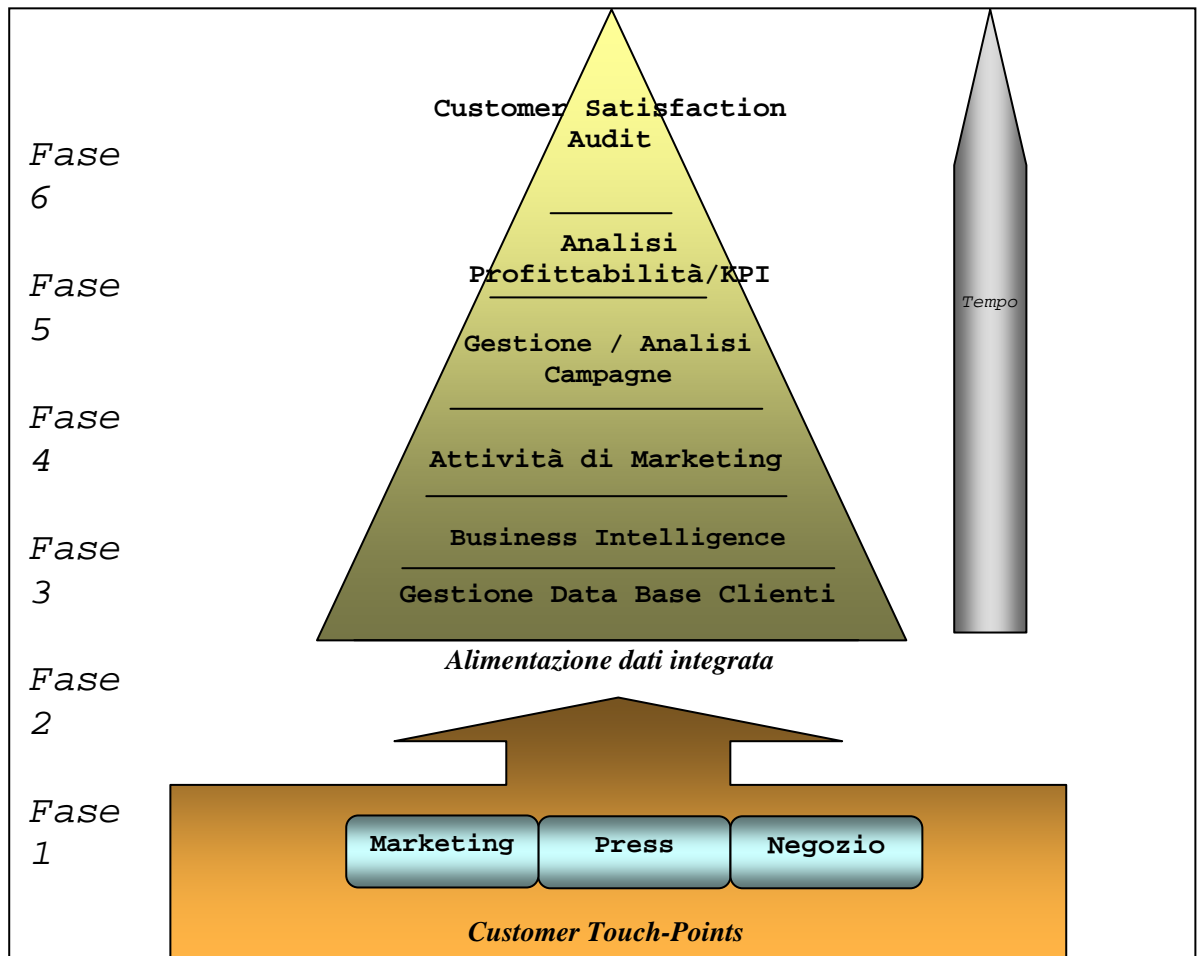
Il caso di studio qui analizzato tratta quindi del primo tentativo di un'azienda di mettere in atto una strategia di CRM, e al di là dell'interesse sulle scelte tecniche di implementazione, è interessante vedere come il processo di adattamento si evolve in modo incrementale ed interattivo.

Con il progetto è nato anche un team CRM inserito nell'area Marketing e composto dal manager di progetto, da un assistente del manager con compiti operativi e da un'analista; l'ufficio interagisce con le altre aree coinvolte direttamente nel progetto, l'area *Retail* (negozi di proprietà del gruppo), l'area *Press e Promozioni* e con i Sistemi Informativi e Organizzazione.

Le responsabilità del team di progetto sono:

- Coordinamento della strategia di relazione cliente del gruppo
- Gestione e analisi del database clienti

- Proposte e collaborazione con l'ufficio Press e Promozioni per attività di marketing varie



**Figura 9: il progetto CRM per fasi critiche**

Le fasi e gli obiettivi individuati dal manager e disposti cronologicamente nella figura si implementano nelle quattro attività proposte nel manuale di M.J.A Berry e G.S. Linoff e introdotte nel capitolo 2.

### 4.3 Osservare che cosa il cliente sta facendo

La prima fase del progetto ha riguardato la raccolta delle informazioni e dei dati sul cliente sia attraverso il recupero e l'integrazione di tutte le fonti dei dati già presenti in

azienda sia attraverso la progettazione e lo sviluppo di un programma per l'automazione e la gestione del contatto con il cliente.

Le fonti di informazione sul cliente presenti in azienda sono i negozi e l'ufficio Press, i negozi dispongono di liste di nominativi di clienti con informazioni anagrafiche annesse basate sulla conoscenza diretta del cliente, alcune di queste informazioni sono su supporti cartacei o elettronici, altre sono nelle teste dei venditori.

L'ufficio *Press*/Promozioni dispone di un *database*, anch'esso composto da informazioni di vario genere e su supporti diversi che riguarda principalmente classi di clienti particolari come i VIP, ecc.

L'obiettivo è quello di accedere a queste informazioni, spesso gelosamente custodite, integrarle per creare una unica base di dati condivisa e visibile da tutti i negozi e gli uffici della sede centrale.

L'altro obiettivo è di essere in grado di dare un nome alle transazioni anonime che avvengono tra il cliente e l'azienda.

Le transazioni oggetto di interesse sono quelle relative all'acquisto del cliente in negozio, tali transazioni sono già automatizzate e memorizzate per motivi fiscali ma sono anonime. Ogni transazione è identificata dalla concatenazione di una serie di attributi come il negozio, la cassa, la data e l'ora a cui è stato effettuata, l'ammontare speso, il tipo di pagamento ecc., ogni transazione (o scontrino) riporta un riferimento ai prodotti acquistati.

Il concetto di associare un nome allo scontrino non è del tutto nuovo in negozio, esiste già un meccanismo di identificazione dello scontrino usato più comunemente per l'emissione di una fattura ma in alcuni casi anche per l'identificazione di una persona. Esiste già, al momento del pagamento alla cassa, la possibilità di associare ad ogni scontrino emesso un codice identificativo di una persona e per ogni codice identificativo può essere compilato un profilo anagrafico. Tali informazioni vengono già raccolte e memorizzate nel database operativo aziendale e sono accessibili.

Gli obiettivi in questo contesto sono quindi di promuovere dalla cassa l'abbinamento dello scontrino ad un profilo cliente e la raccolta di informazioni sul cliente.

La procedura si articola nelle seguenti fasi; al momento del pagamento viene chiesto al cliente se ha già rilasciato i suoi dati prima in un qualsiasi negozio di proprietà GA, in caso affermativo si recupera l'anagrafica di quel cliente già inserita nel DB e si abbina,

attraverso un meccanismo automatizzato, lo scontrino effettuato a quel cliente. Nel caso in cui il cliente non risultasse già presente nel DB gli viene chiesto di compilare un modulo cartaceo con i dati personali di base e di firmare per rilasciare il consenso al trattamento dei propri dati, i dati così raccolti vengono inseriti nel sistema, viene creato un nuovo profilo anagrafico a cui finalmente è abbinato lo scontrino appena effettuato.

Dato che la procedura non è interamente automatizzata e data la sua complessità il meccanismo è suscettibile di errore. La qualità dei dati così raccolti presenta due problemi principali, errore di compilazione delle informazioni del profilo anagrafico o mancata associazione dello scontrino al codice cliente.

Il primo di tipo di errore può essere recuperato con il tempo, con l'esperienza del personale di cassa e con la manutenzione continua dei dati. Il secondo tipo di errore è più grave, perché ogni scontrino emesso e non associato ad un cliente è un'informazione persa e praticamente impossibile da recuperare.

Inoltre trattandosi di dati personali si esige un comportamento da parte dell'azienda in totale accordo con il codice sulla tutela dei dati personali (che sostituisce ed integra la legge 675/96), che rende ancor più delicato e complesso il processo di raccolta e archiviazione del dato.

Per incentivare la raccolta dei dati è stato concordato con i manager *retail* che il raggiungimento di determinate soglie nel tasso di associazione cliente-scontrino diventasse un obiettivo di negozio.

I dati anagrafici di base che vengono chiesti durante la compilazione del modulo privacy sono: Titolo, Nome, Cognome, Indirizzo, Città, Provincia, Nazionalità, Data di Nascita, e-mail, telefono, età (range).

Lo sviluppo del programma di gestione dei clienti è stato affidato ad una società di consulenza esterna, che ha agito sotto le specifiche del manager di progetto di CRM, ed è tuttora in via di sviluppo; il programma poggia sulla base di dati operativa dell'azienda e gestisce in particolar modo le tabelle dove si trovano i dati anagrafici relativi ai clienti. Il programma funge da interfaccia tra la base di dati (dei clienti) e gli utenti; esistono due tipi di utenti che accedono ai dati con privilegi diversi e con finalità diverse, i negozi e l'ufficio CRM e *Press/Promozioni*.

Il programma ha due funzioni principali: di gestione della base di dati del cliente e di gestione delle campagne marketing.

I negozi usano il programma, per quanto riguarda la gestione del *database*, per abbinare un profilo anagrafico allo scontrino, per inserire nuovi profili anagrafici o modificarli, stampare *report* di controllo sui tassi di associazione del negozio o di performance dei venditori. Per quanto riguarda invece la parte di gestione delle campagne marketing, i negozi possono controllare dal programma le campagne marketing in corso e le loro caratteristiche e stampare le liste di clienti coinvolti in una campagna.

L'ufficio CRM può dalla sede centrale inserire o modificare sia manualmente sia automaticamente, attraverso l'acquisizione di liste di nominativi, le anagrafiche clienti, dispone inoltre di funzionalità di manutenzione dei dati come ad esempio la funzione di ricerca e cancellazione di codici diversi associati alla stessa persona. L'ufficio CRM è l'interfaccia tra i negozi e gli sviluppatori del sistema di gestione *data base* clienti ed è di supporto per l'utilizzo di tutte le funzionalità del programma.

L'ufficio CRM gestisce le campagne marketing in tutte le fasi di vita della campagna, dall'ideazione alla pianificazione dalla messa in atto alla misurazione dei risultati, il *software* di gestione delle campagne permette di creare, acquisire ed elaborare liste di clienti da inserire in una campagna, ogni campagna è codificata in maniera univoca e memorizzata nel sistema con i codici dei clienti coinvolti e i criteri utilizzati per la scelta di quel target, mentre dal lato del DB del cliente, ogni cliente ha associate le campagne di cui è stato oggetto.

#### **4.3.1 Azioni da intraprendere**

I dati raccolti fino a questo momento dall'azienda provengono essenzialmente da un unico canale, il negozio, esistono però altri punti di contatto tra l'azienda e il cliente rappresentati dal sito web, la posta tradizionale e quella elettronica e il *call center*. Per il momento non è possibile attraverso nessuno di questi canali effettuare degli acquisti, si tratta di contatti di tipo informativo che potrebbero andare a formare un apparato di assistenza e informazione verso i propri clienti.

Per alcuni di questi canali esiste un meccanismo di memorizzazione della transazione ma manca l'identificazione della persona interagente. I dati provenienti dalla memorizzazione dei contatti avvenuti tramite questi canali dovrebbero essere integrati con il programma di gestione del *database* del cliente. Il sistema di assegnazione di un

codice ad un nuovo cliente o *prospect*, congiuntamente alla richiesta delle informazioni anagrafiche di base e il consenso al trattamento dei dati personali dovrebbe riguardare anche questo nuovo modo di interagire con il cliente, dovrebbe poi essere predisposta una modalità di memorizzazione del contenuto e della tipologia di transazione avvenuta con il cliente. Ad esempio dovrebbe essere possibile distinguere tra un reclamo e una richiesta di informazione su un prodotto e dovrebbe essere possibile approfondire il motivo del reclamo o la tipologia di informazione richiesta.

Visto che il cliente potrà interagire da più punti con l'azienda è necessario scegliere una procedura standard di relazione con il cliente, indipendentemente dal canale infatti l'azienda deve apparire all'esterno come un tutto unico e coerente.

La seconda azione da intraprendere è di continuare il processo di automazione delle campagne marketing e sofisticazione del programma di gestione.

Il processo di automazione dovrebbe comprendere anche la gestione delle risposte alle campagne e funzionalità di misurazione dei risultati, la misurazione dei risultati di una campagna in un contesto non di marketing diretto può essere molto difficile ma è indispensabile non solo per motivi di controllo ma anche per motivi di analisi, l'aspetto dell'analisi verrà discussa a fondo nel capitolo successivo.

Il programma dovrebbe anche offrire strumenti di segmentazione per la selezione dei clienti da inserire in una campagna, se la segmentazione oltre a basarsi su i dati anagrafici e di comportamento di acquisto si baserà anche sui dati di ritorno delle precedenti campagne la segmentazione acquisterà anche un valore predittivo sul tipo di risposta che il cliente potrebbe dare ad un certo tipo di campagna.

#### **4.4 Ricordare che cosa i clienti hanno fatto nel passato**

Il primo disegno del *datawarehouse* del cliente è avvenuto sotto le specifiche del manager di progetto del CRM che ha ipotizzato e cercato di anticipare le possibili esigenze di analisi a cui il *datawarehouse* doveva rispondere. Con l'esperienza di analisi alcune caratteristiche sia tecniche che di contenuto (aggiunta di nuovi attributi) sono state modificate.

Il *datawarehouse* è composto da dati provenienti da diverse fonti, il DB del venduto, del prodotto e del cliente. Nel DB del venduto sono archiviati tutti gli scontrini emessi

nei negozi di proprietà con tutte le informazioni che riguardano la vendita come il negozio, la data, il tipo di pagamento, l'ammontare speso, un riferimento ai prodotti acquistati e il cliente che ha effettuato l'acquisto (nel caso lo scontrino sia stato abbinato ad un cliente) ecc.. Il prodotto ha un'organizzazione piuttosto complessa che rispecchia la realtà del business della moda, il prodotto è identificato da una gerarchia di concetti che ne specifica le caratteristiche, la banca dati del prodotto è condivisa in tutta l'azienda, sia nel contesto di vendita al consumatore finale nei negozi di proprietà, nella vendita a negozi non di proprietà sia nel *whole sale* (negozi multimarca).

Il *datawarehouse* deve rappresentare i dati e le informazioni sui clienti così come li vedono i manager ma deve mantenere il contatto con la fonte dei dati e quindi con le modalità e la terminologia con cui vengono presentati dal programma di gestione del *database* clienti.

In questa fase pilota il *datawarehouse* per il CRM è stato implementato con delle viste SQL che estraggono i dati direttamente dalla base di dati operativa acceduta da tutti gli utenti dell'azienda con richieste tipicamente operative; il motivo di tale scelta implementativa è da imputarsi a ragioni di costo e ragioni di flessibilità.

La granularità del dato è al livello di dettaglio dello scontrino ma l'informazione del prodotto viene mantenuta e gestita come dimensione per cui sono recuperabili anche le informazioni a livello di riga di scontrino, i dati sono forniti all'utente perché questo possa aggregarli, vederli da più punti di vista, navigarne le gerarchie.

Nella figura 10 è riportato un esempio semplificato del *datawarehouse*, per motivi di semplicità lo schema ha valenza solo esemplificativa.

Il *datawarehouse* presenta uno schema a fiocco di neve: la tabella dei fatti (VENDITA-CLIENTE) è composta dalle chiavi che agganciano le tabelle delle dimensioni (CLIENTE, NEGOZIO, PRODOTTO), nella tabella dei fatti ci sono inoltre tutti gli attributi misura che caratterizzano il singolo scontrino come *Ammontare*, e tutte le misure calcolate utili per l'analisi come *Q.tà capi*.

Dal momento che il *datawarehouse* è implementato con delle viste non materializzate sulla base di dati operativa non esiste un vero e proprio *datawarehouse* del cliente in azienda ma una sorta di simulatore di esso che presenta i dati all'utente organizzati secondo una logica multidimensionale. Questo tipo di organizzazione implica che ogni interrogazione sia un'interrogazione alla base di dati operativa e che l'analista debba

ogni volta elaborare i dati ottenuti dal massimo dettaglio fino al livello di sintesi desiderato.

Questa soluzione rappresenta quella che dà il massimo grado di flessibilità in termini di modificabilità ed estendibilità del sistema di volta in volta che vengono realizzate nuove funzioni nella parte operativa, nella fase iniziale di un progetto la flessibilità è la caratteristica più importante perché permette la modifica dinamica e interattiva del *datawarehouse* a seconda delle esigenze che emergono dall'analisi e dei cambiamenti continui del programma di gestione del *database* del cliente.

Sviluppare un *datawarehouse* con un ambiente dedicato è un'operazione che richiede un investimento ingente sia da un punto di vista di costi sia da un punto di vista di tempo (minimo 6 mesi-uomo), si è scelto quindi una soluzione molto meno costosa ma che fungesse da soluzione pilota per testare l'effettiva esigenza e utilità di un *datawarehouse* per il CRM. La flessibilità della soluzione scelta ha permesso inoltre di modificare facilmente e in breve tempo le caratteristiche del *datawarehouse* in base alle specifiche richieste di volta in volta dall'analista.

Gli svantaggi della soluzione scelta riguardano essenzialmente il fatto di non operare su un ambiente dedicato all'analisi multidimensionale, cioè su una copia storica del DB ma di agire in concorrenza con tutti gli altri utenti della base di dati transazionale con il conseguente decadimento delle prestazioni per quanto riguarda i tempi di elaborazione ed estrazione dei dati. Il decadimento delle prestazioni è reso drammatico dal modo in cui si vanno ad interrogare i dati, le interrogazioni tipiche per le analisi multidimensionali interrogano centinaia di migliaia di record e per fornire i dati aggregati sulle varie dimensioni vengono implicate join molto costose.

Quando il programma di gestione del database clienti è stato ampliato con il modulo di gestione delle campagne marketing il *datawarehouse* ha subito le conseguenti modifiche. In figura 10 nello schema a fiocco di neve del *datawarehouse* è rappresentata la tabella, aggiunta successivamente rispetto alle altre, per le campagne marketing.



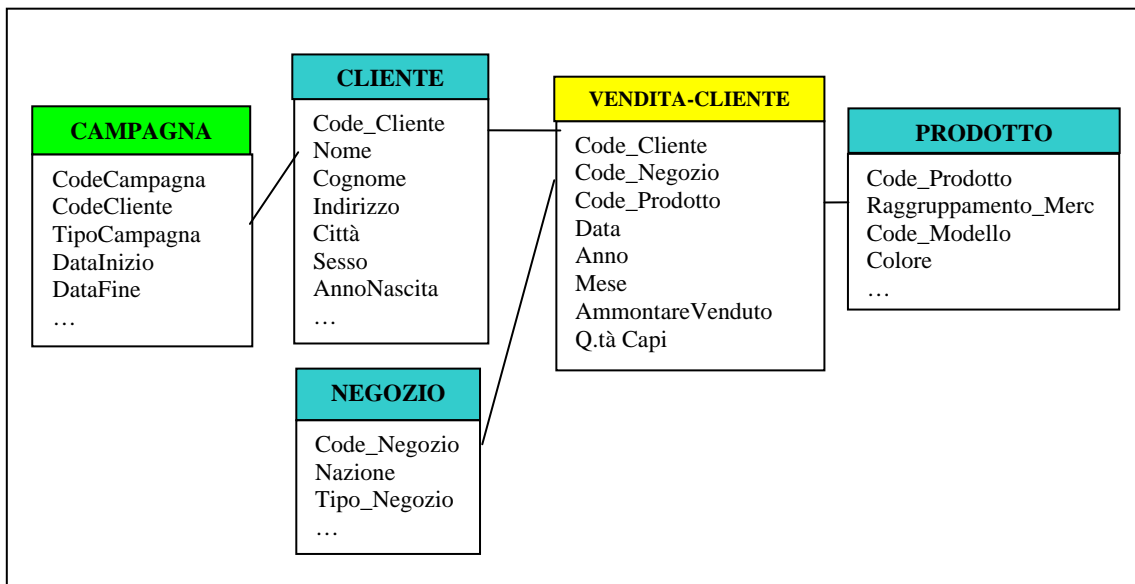


Figura 10: Lo schema del *datawarehouse* ampliato per le campagne

#### 4.4.1 Azioni da Intraprendere

Una volta ricevuto il consenso sull'intero progetto si penserà a portare il *datawarehouse* del CRM in un ambiente dedicato dove verranno curati anche gli aspetti tecnici di performance di accesso ed elaborazione dei dati attraverso l'uso di indici, viste materializzate, ecc.

Dal momento che si prevedono dei cambiamenti nel sistema per il momento è appropriato che la struttura del *datawarehouse* rimanga ancora così flessibile. I cambiamenti coinvolgono la gestione dei canali, l'ampliamento delle funzionalità del programma di gestione del *database* clienti e delle campagne marketing e l'elaborazione degli indicatori.

Una volta che le altre fonti di dati sul cliente saranno integrate alla base di dati operativa anche il *datawarehouse* dovrà subire le modifiche necessarie per riportare i nuovi fatti nell'ambiente di analisi.

Dalle analisi condotte sui dati del *datawarehouse* l'analista e i manager hanno individuato una serie di indicatori, calcolati sui dati, di un certo interesse e che vengono usati molto spesso nelle analisi; il *datawarehouse* dovrebbe fornire quegli indicatori come un oggetto precalcolato. Nelle prossime fasi vedremo come si giunge all'individuazione degli indicatori.

Quello descritto fin qui è un approccio progettuale totalmente incrementale, come spesso accade nel caso di progettazione del *datawarehouse*, in questo caso però è ancora più accentuato perché ci troviamo nella fase iniziale dell'intero sistema.

#### **4.5 Imparare da ciò che ci ricordiamo**

Una volta che i dati possono essere acceduti con logiche multidimensionali, a questo livello non è importante quale sia l'implementazione del *datawarehouse*, si può procedere con l'analisi.

Dal momento che la prima versione dell'universo CRM non era stata testata da alcuno le prime analisi hanno avuto l'obiettivo di testare la correttezza, l'usabilità e la conformità delle informazioni rispetto ai fatti che si vogliono analizzare.

Il test della correttezza consiste nel controllo di eventuali errori nei dati estratti dal *datawarehouse* rispetto ai dati originari nel DB, causati da logiche errate di estrazione del dato dal DB e calcolo di questo per creare il corrispondente oggetto del *datawarehouse*.

In questa fase è emersa l'importanza del collegamento e della comunicazione tra il team di sviluppo del programma di gestione e quello di sviluppo del *datawarehouse*.

La seconda fase di test ha riguardato la conformità tra l'universo e i fatti da analizzare; l'universo è stato progettato sotto le specifiche del manager di progetto prima che alcuni ampliamenti importanti fossero apportati al programma di gestione clienti ( ad esempio il modulo gestione campagne), le indicazioni del manager tracciavano una prima approssimazione delle caratteristiche dell'universo e delle esigenze analitiche a cui l'universo doveva rispondere.

E' stata necessaria un'approfondita esplorazione dei dati per valutare l'eventuale mancanza di informazioni nel *datawarehouse* o la necessità di aggiungere degli oggetti calcolati a partire da quelli esistenti e che si ritenevano particolarmente utili per l'analisi.

L'esplorazione dei dati si è accompagnata alla conoscenza degli strumenti di interrogazione ed elaborazione di *report*, *BusinessObjects*; è stato possibile una prima valutazione sulla tipologia di strumento, sulle sue potenzialità e facilità d'uso.

Lo strumento di interrogazione dei dati, elaborazione e *reporting* a disposizione dell'analista è *BusinessObjects-BusinessObjects*.

Sul datawarehouse è stato disegnato l'universo del CRM con lo strumento *BusinessObjects-Designer*, un universo *BusinessObjects* funge da interfaccia tra i dati e l'utente, presentando i dati organizzati secondo le logiche e la terminologia conosciuta dall'azienda e in maniera trasparente rispetto alle problematiche tecniche del database.

I nomi delle variabili sono gli stessi nomi usati nel programma di gestione del CRM e la presentazione grafica dell'universo all'utente è costituita dalla lista di variabili suddivise in cartelle, cartella dell'*Anagrafica Cliente*, del *Venduto*, delle *Campagne*, ecc.

La funzionalità di interrogazione dei dati è accessibile dal pannello delle query come mostrato in figura 11, l'interfaccia permette di creare la propria interrogazione selezionando gli oggetti dalla lista delle variabili a sinistra. E' possibile visualizzare l'SQL generato corrispondente alla query.

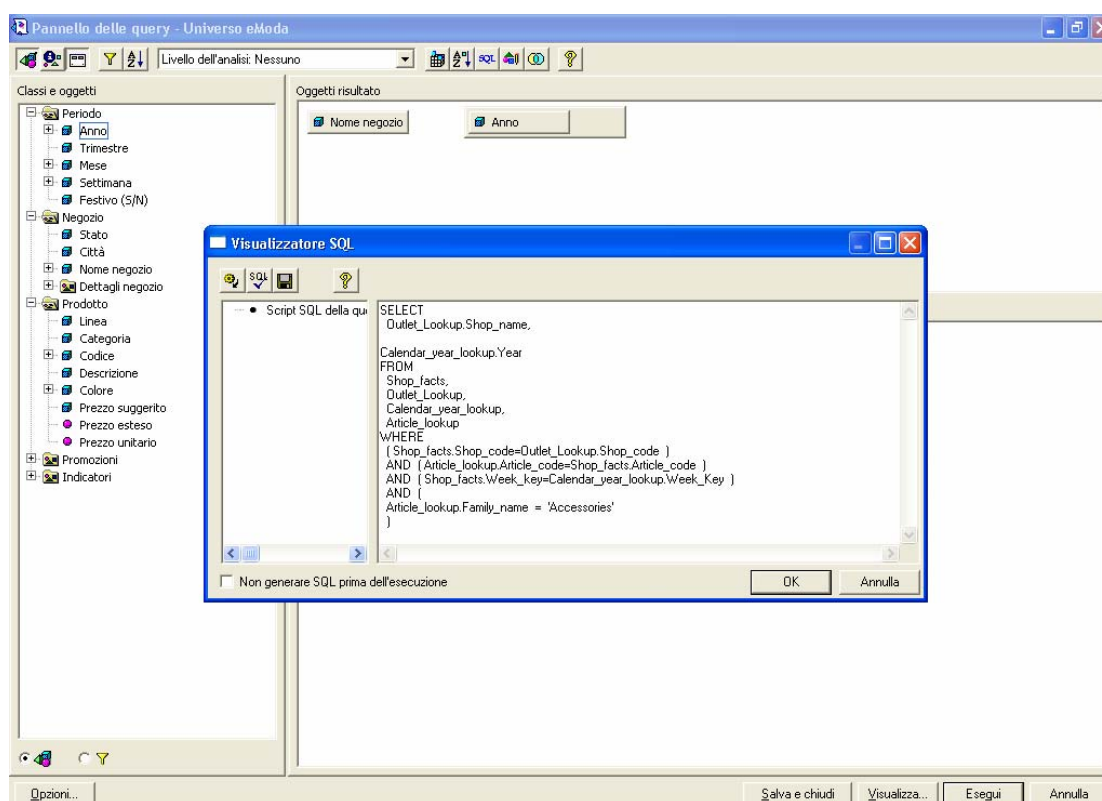


Figura 11: Esempio di pannello delle query fornito da BusinessObjects

Una volta chiarita la situazione allo stato dell'arte sia dal punto di vista dei dati disponibili, delle tecnologie e strumenti da usare e i team e i ruoli con cui l'analista si interfaccia, è stato necessario che l'analista e il manager di progetto ritornassero sulla definizione degli obiettivi dell'analisi sul *datawarehouse*, sulle aspettative riguardo al tipo di informazioni restituite dall'analisi e sull'utilità di esse.

#### 4.5.1 L'analisi dei dati

Esaurita la fase principale di test e risolti i problemi emersi da tale test è iniziata la prima analisi sul contenuto del DB; il focus dell'analisi è stato inizialmente l'anagrafica del cliente dal momento che tali informazioni rappresentano la vera novità apportata dal CRM.

Mentre infatti i dati di vendita sono ovviamente già ampiamente studiati ed analizzati, anche se in un'ottica diversa da quella orientata al cliente, i dati sulle reali caratteristiche anagrafiche provenienti direttamente dai clienti suscitavano un certo interesse e curiosità.

Il tipo di analisi effettuate è pertanto di statistica descrittiva, per calcolare la distribuzione sull'intera popolazione dei clienti, o su sottogruppi di clienti, in base a variabili come la nazionalità, l'età, il sesso, ecc.

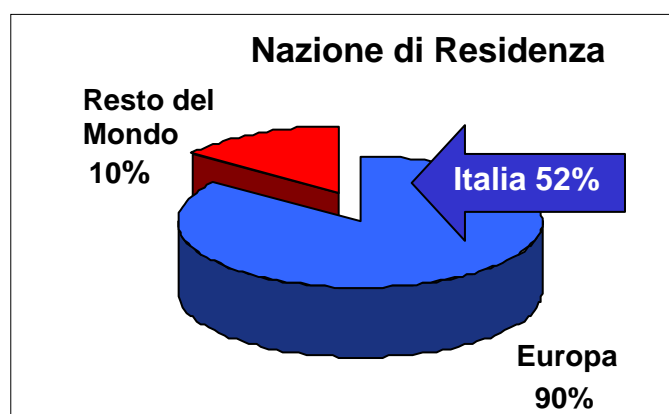


Figura 12: Esempio di analisi descrittiva

Le premesse già fatte sulla qualità dei dati anticipano il fatto che le informazioni contenute nel DB non hanno potuto soddisfare a pieno le curiosità né è stato possibile utilizzarli in analisi più sofisticate di quelle meramente descrittive.

Le modalità di reperimento delle informazioni sul cliente fino ad ora utilizzate, e cioè l'intervista direttamente al cliente attraverso il riempimento del modulo privacy, non permettono di approfondire la conoscenza del cliente su aspetti socio-demografici molto interessanti quali la famiglia, gli interessi personali, il tenore di vita. Al momento le informazioni anagrafiche disponibili sono l'età e il sesso oltre all'indirizzo, il comune e il paese di residenza, ma anche i risultati che si possono ottenere con queste poche informazioni possono essere di grande importanza e stimolare il management ad approfondire la conoscenza del proprio cliente innanzitutto attraverso l'attivazione del monitoraggio dei clienti anche da altri canali come il sito web oppure attraverso altre attività come ad esempio interviste approfondite e *focus group* e attraverso l'acquisto di informazioni sui propri clienti da altre banche dati.

Esplorato l'universo dei dati e prodotte una serie di fotografie, più o meno dettagliate e articolate, del tipo di cliente con cui l'azienda intrattiene rapporti la seconda fase di analisi descrittiva si è incentrata sull'analisi del venduto dal punto di vista del comportamento di acquisto del cliente, l'oggetto di osservazione quindi è il cliente. Per comportamento di acquisto si intende l'insieme delle caratteristiche che contraddistinguono gli acquisti di un cliente, dal numero di scontrini effettuati in un certo intervallo di tempo, al numero di negozi diversi visitati, al prodotto più frequentemente acquistato, ecc.; la novità rispetto agli studi sul venduto è che tracciando gli acquisti del cliente è possibile analizzarne il comportamento, l'oggetto osservato non è più solo uno scontrino anonimo ma un gruppo di scontrini acquistati dalla stessa persona.

La scelta di quali caratteristiche del comportamento mettere in evidenza è stata concordata tra l'analista e il manager di progetto in base a un criterio di rilevanza e significatività dell'informazione ma soprattutto in base all'esistenza di indicatori già utilizzati in azienda in studi precedenti sul cliente e quindi facilmente comprensibili e confrontabili.

Il comportamento di acquisto di un cliente è descrivibile da molti punti di vista, l'Azienda considera in maniera piuttosto distinta i clienti di ogni insegna o tipologia di

negozio, la prima dimensione di analisi da approfondire è stata infatti il negozio e i risultati delle analisi hanno mantenuto la distinzione per tipologia di negozio concentrandosi sulla descrizione del comportamento di acquisto dei clienti che acquistano in ogni diversa tipologia.

Il comportamento di acquisto si può descrivere con indicatori che misurano le caratteristiche degli acquisti di ogni cliente come quanti scontrini effettua nello stesso negozio all'anno (frequenza), come varia la distribuzione dell'ammontare speso durante i vari acquisti (*up-selling*), quanti clienti frequentano tipologie di negozi diversi (*cross-selling*) e quanti clienti ripetono l'acquisto il periodo di tempo successivo o abbandonano la marca (*retention*).

In figura 13 è riportato il grafico che mostra la suddivisione in quattro gruppi di clienti (asse delle ascisse) in base all'ammontare speso e sull'asse delle ordinate quanto ogni gruppo ha generato sul totale acquistato in ogni tipologia di negozio (X, Y, Z,...). Questo tipo di descrizione è molto interessante e serve per vedere quanto pesano, quanto sono importanti, ogni gruppo di clienti (il primo 5%, il secondo 15%, ecc.) rispetto agli altri e rispetto al totale acquistato.

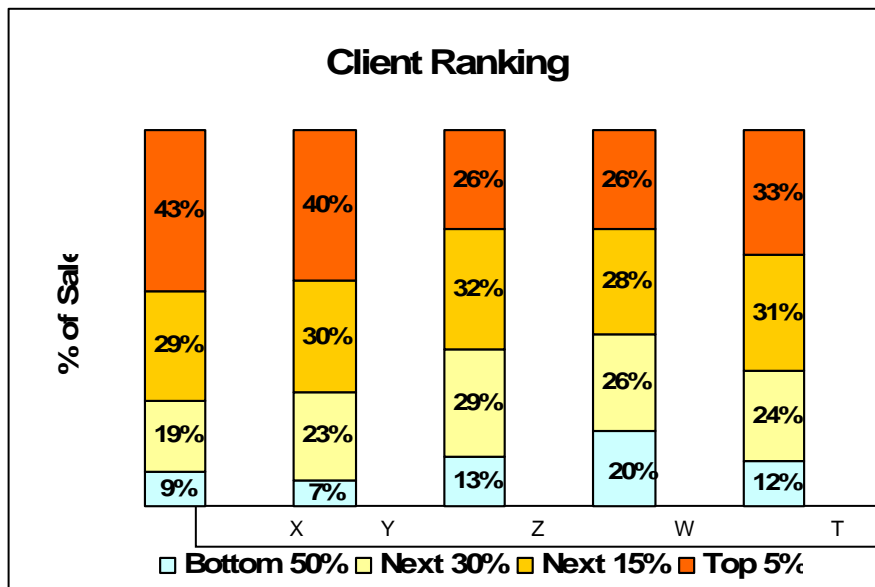


Figura 13: Esempio di analisi di comportamento di acquisto

#### 4.5.2 Strumento di analisi

Lo strumento utilizzato sia per l'interrogazione dei dati sia per l'analisi multidimensionale è *BusinessObjects*, mentre per la formattazione dei report si è deciso di utilizzare sempre un foglio elettronico tipo excel, il risultato delle analisi su *BusinessObjects* in forma tabellare viene riportato su un file excel utilizzando le funzionalità grafiche e di presentazione tipiche del foglio elettronico e già ben conosciute in azienda.

*BusinessObjects* si è rivelato uno strumento di analisi piuttosto flessibile, è possibile creare nuove variabili che rimangono locali al documento, effettuare calcoli sulle variabili ed applicare funzioni di tipo numerico, di carattere, logico e funzioni di aggregazione.

Lo strumento distingue essenzialmente due tipi di oggetti, l'attributo tipo dimensione e gli oggetti misura che sono oggetti calcolati su cui è definita una funzione di aggregazione. Nell'esempio di Figura 6, *AmmontareVenduto* è una misura su cui è definita come funzione di aggregazione la somma, se creo una tabella di due colonne formata da una dimensione e una misura, poniamo *Code\_Negozio* e *Ammontare*, la misura darà come risultato per ogni riga (per ogni negozio) l'ammontare incassato da quel negozio.

*BusinessObjects* permette di visualizzare i risultati di una query in forma tabellare, in forma tipo tabella-pivot e come grafico (a torta, a barre, ecc.), *BusinessObjects* inoltre permette di integrare in un unico documento i risultati di più di un'interrogazione o di query eseguite su differenti sorgenti di dati.

*BusinessObjects* è uno strumento di analisi molto flessibile, pur non offrendo delle funzioni di analisi statistica come il *rank* o il calcolo di frequenze che si adattano facilmente al contesto multidimensionale è possibile, con un po' di inventiva, ottenere i risultati desiderati.

In alcuni casi però le funzionalità disponibili all'utente non sono bastate per lo scopo di analisi ed è stato necessario chiedere al *designer* dell'universo di creare appositamente un oggetto che risolvesse il problema o di creare una query SQL ad hoc.

I casi che non potevano essere risolti facilmente con *BusinessObjects* sono casi dove tipicamente il contesto di riferimento deve essere calcolato in maniera molto complessa

e con *BusinessObjects* tale complessità diventa molto difficile se non impossibile da gestire.

In generale, indipendentemente dallo strumento usato per l'analisi i dati organizzati in modo multidimensionale non permettono con una sola *query* di ottenere indicatori aggregati relativamente a gruppi separati e indipendenti di dimensioni. Segue un esempio di analisi, il cui risultato è riportato in figura 14 la frequenza di acquisto, in cui ci si trova con tale problematica dei dati nelle due tabelle.

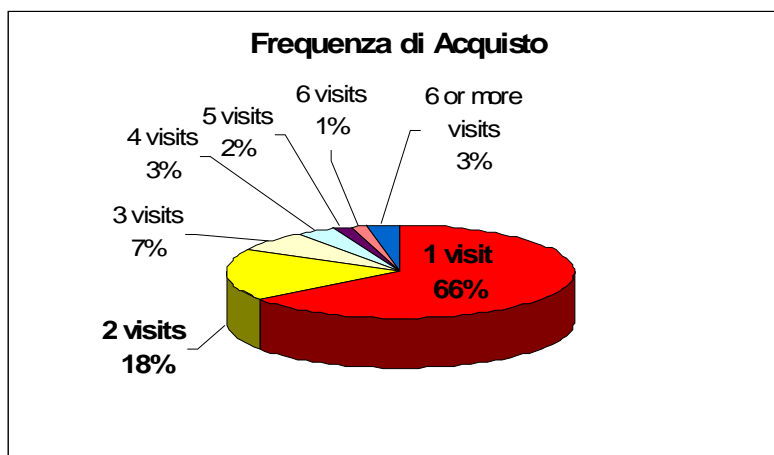


Figura 14: Esempio di analisi del comportamento d'acquisto

Il significato del grafico riportato in figura 10 è che sulla totalità dei clienti che hanno acquistato almeno uno scontrino la percentuale dei clienti che hanno fatto un solo scontrino (nel periodo di tempo considerato) è il 66%, i clienti che hanno fatto 2 scontrini rappresentano il 18%, ecc..

Per arrivare a trovare la frequenza di acquisto si genera una *query* sul *datawarehouse* con *CodeCliente*, *CodeScontrino* e il risultato è quello esemplificato dalla tabella che segue.

CodeCliente	CodeScontrino
Cliente1	Sc1
Cliente2	Sc1
Cliente2	Sc2

Per calcolare la frequenza di acquisto è necessario calcolare il numero di scontrini che ogni cliente ha effettuato e si raggruppa su cliente. Il risultato è quello esemplificato nella tabella seguente.



<b>CodeCliente</b>	<b>Count(distinct CodeScontrino)</b>
Cliente1	1
Cliente2	2

A questo punto sarebbe necessario contare il numero di clienti che hanno fatto lo stesso numero di scontrini e raggruppare sul numero di scontrini ma questo non è possibile perché non esiste una dimensione numero di scontrini e il risultato della Count(distinct CodeScontrino) non è una dimensione sulla quale aggregare.

Per risolvere tale problematica dei dati con *BusinessObjects* è possibile creare degli oggetti pre-calcolati con *BusinessObjects-Designer*, oppure fare una prima query in cui si esegue il primo conteggio su una dimensione e poi eseguire una seconda query con la seconda dimensione non aggregata, aggregare sulla prima dimensione e integrare i risultati delle due *query* in un unico documento.

Nel caso dell'elaborazione per la *Frequenza di Acquisto* il *Designer* dell'universo ha aggiunto un oggetto che conta il numero di scontrini e che può essere usato come dimensione in *query* che escludono lo scontrino e dimensioni di dettaglio dello scontrino come il prodotto (righe di scontrino).

#### **4.5.3 Analisi per la gestione delle campagne marketing**

L'analisi dei dati per la produzione di informazioni e conoscenza per il supporto della gestione delle campagne merita una trattazione separata perché, anche se la tecnologia e le tecniche di analisi non si discostano di molto da quelle precedentemente discusse, la gestione delle campagne ha un carattere non solo strategico ma anche operativo, attraverso le campagne infatti si mettono in pratica le decisioni prese, la campagna è il punto di contatto tra le decisioni dell'azienda e il cliente.

L'analisi ha due obiettivi principali, uno di definizione del target verso cui indirizzare una campagna, l'altro è di misurazione dei risultati dell'attività di marketing e della sua efficacia.

Con la definizione del target si stabiliscono i criteri di segmentazione con cui i clienti sono selezionati per una certa campagna (ad esempio, clienti con età sotto i 25 anni, solo donne, solo residenti a Milano, ecc.) e questa selezione è attuabile dal programma

di gestione campagne, spesso però è necessario usare criteri più sofisticati o criteri multipli che il programma ad oggi non supporta.

Per superare questa rigidità del sistema nel caso di criteri particolarmente complessi la selezione del target avviene elaborando i dati con *BusinessObjects*, e infine creando una lista di clienti acquisibile dal sistema di gestione campagne.

Per quanto riguarda la scelta strategica dei destinatari cui indirizzare una specifica campagna marketing la decisione finale è presa dal manager di progetto, in generale però il criterio più comunemente utilizzato è l'ammontare di spesa effettuato dal cliente in un certo periodo di tempo, il primo raggruppamento della popolazione dei clienti, il più immediato, è per fasce di spesa.

Raggruppare i propri clienti per fascia di spesa è il criterio più immediato ma ad un'attenta analisi può non essere quello che meglio rappresenta il grado di profittabilità del cliente per l'azienda.

Visto il tipo di dati presente in azienda è stato scelto un metodo che si basasse sulle variabili del comportamento di acquisto, nella fattispecie il metodo *RFM* (*Recency, Frequency, Monetary Value*). Il metodo RFM è il più diffuso in letteratura come punto di partenza per le aziende che vogliono applicare tecniche di segmentazione sui propri clienti.

Il metodo prende in considerazione tre aspetti del comportamento di acquisto del cliente, la *Recency*, che misura quanto è recente l'ultimo acquisto del cliente, la *Frequency*, il numero di volte che il cliente ha acquistato in un certo periodo di tempo e infine il *Monetary Value* che indica la spesa effettuata in un certo periodo di tempo.

Una volta misurati questi comportamenti di acquisto si valuta l'importanza che ognuno dei tre comportamenti ha per l'azienda, ad esempio si può valutare che sia più importante che un cliente acquisti spesso invece che faccia una spesa molto consistente ma una sola volta; decidere il criterio di importanza tra i tre indicatori è l'operazione più delicata, in letteratura l'indicatore considerato più importante è tradizionalmente la *Recency*.

Ogni cliente avrà un valore RFM, dato dalla somma pesata (secondo il criterio di importanza scelto) dei valori raggiunti per ognuno dei tre indicatori, il risultato finale sarà la suddivisione dei clienti in quattro gruppi diversi.

Il criterio RFM può essere arricchito con altri attributi sia quantitativi che qualitativi, il fine principale dell'applicazione di criteri come la RFM è quello di raggruppare i clienti simili in modo tale da rendere la gestione dei clienti più snella ed efficiente ma anche quella di diffondere tra chi è a contatto con il cliente una visione comune e condivisa di chi è il cliente e quale è il suo valore per l'azienda. Il metodo non serve per segmentare la popolazione dei clienti ma è un metodo di *scoring* che serve cioè ad attribuire un punteggio ad ogni cliente in base al valore che questo rappresenta per l'azienda.

Invece la segmentazione di un insieme di clienti consiste nella suddivisione dell'insieme in sottoinsiemi (segmenti) spesso ma non necessariamente disgiunti che presentino caratteri di omogeneità al loro interno e disomogeneità rispetto agli elementi appartenenti a segmenti diversi.

Una volta che sono stati individuati i segmenti, l'impresa elabora una strategia per "agire" su tutti o su alcuni di questi segmenti; per ogni segmento si decideranno azioni di tipo diverso, calibrate sulle caratteristiche dei clienti di quel segmento. Nel capitolo successivo verrà discussa la questione in modo approfondito.

La misurazione dei risultati in termini quantitativi è un passo necessario per poter compiere una valutazione dell'efficacia e dell'efficienza (in termini di costi, ad esempio costo per contatto) della campagna marketing e del comportamento del cliente rispetto alle campagne.

Nel caso dell'Azienda in questione non si parla di marketing diretto ma per campagne marketing si intende tutte quelle attività di contatto esplicito tra l'azienda e il cliente, le tipologie di campagne sono molto diverse, gli obiettivi della campagna possono essere diversi e ogni campagna può avere più obiettivi, ogni obiettivo può essere implementato con una varietà di tecniche:

- Promozione di un prodotto (lancio pubblicitario sui media, concorsi a premi, eventi di lancio, spedizione catalogo collezione)
- Promozione dell'immagine dell'azienda (Pubblicità sui media, testimonial, eventi)
- Acquisizione di nuovi clienti (eventi, lotterie)
- Fidelizzazione clienti (spedizione catalogo pubblicitario collezione, eventi, particolari promozioni, carta fidelity)

Alcune campagne si rivolgono ai clienti in maniera indiscriminata, altre a gruppi selezionati di clienti conosciuti, altre ancora a gruppi di persone che l'azienda non riconosce come propri clienti (prospect o suspect).

La misurazione dei risultati di una campagna di marketing non diretta non è immediata, inoltre per la maggior parte degli obiettivi citati sopra è molto difficile determinare e quantificare il grado di raggiungimento dell'obiettivo.

In generale l'indicatore che indica il grado di risposta alla campagna è la *Redemption*, che è un indicatore che esprime sia i risultati in termini di efficacia che di efficienza perché rapporta i risultati ottenuti con il costo che tale attività ha comportato.

La tipologia di risposta o *Redemption* varia in base alla tipologia di campagna e non è sempre di facile individuazione e definizione, come per l'analisi dei dati del *datawarehouse* la scelta degli indicatori più espressivi è il frutto di una fase di studio e riflessione tra l'analista, il manager di progetto CRM e gli altri manager coinvolti a seconda del tipo di campagna.

La *Redemption*, per quelle campagne che prevedono il coinvolgimento di clienti già conosciuti e presenti nel *database*, è calcolata in base al comportamento di acquisto dei clienti coinvolti successivamente al contatto.

Affinché i risultati della campagna siano misurabili i dati riguardo ad essa devono essere memorizzati e organizzati, il programma di gestione campagne memorizza tutte le informazioni relative alla tipologia di campagna e ai clienti coinvolti, il programma però non offre strumenti dedicati per la misurazione dei risultati, l'uso combinato del programma con *BusinessObjects* rende possibile l'analisi del comportamento di acquisto dei clienti coinvolti.

Con *BusinessObjects* vengono estratti solo i clienti della lista della campagna che si sta analizzando (nella *where* condizione che seleziona solo il codice della campagna corrispondente), gli attributi selezionati invece riguardano tutte le informazioni che si vogliono conoscere del cliente e del suo comportamento di acquisto.

Il documento *BusinessObjects* contiene quindi solo e unicamente i dati relativi ai clienti della campagna e questi dati possono venir elaborati come al solito attraverso aggregazioni sul tempo, sul negozio e il prodotto e attraverso l'applicazione di funzioni di statistica descrittiva.

Il risultato delle analisi consiste nella descrizione del comportamento di acquisto dei clienti coinvolti nella campagna.

Dal momento che la tipologia di campagne marketing discussa non permette una verifica puntuale e diretta della risposta è necessario studiare e analizzare le variabili dei comportamenti tracciati per verificare se essi possano essere in qualche modo riferibili all'impatto della campagna o meno. Una corretta procedura di gestione campagne impone che venga sempre scelto un campione di controllo statisticamente significativo, cioè di numerosità adeguata e con le stesse caratteristiche dei clienti coinvolti nella campagna, in secondo luogo impone di ripetere la campagna nel tempo mantenendone le stesse caratteristiche. Queste metodologie di convalida dei risultati misurati sono applicabili sotto certe condizioni: perché venga usato un campione di controllo il numero di clienti nel *database* deve essere sufficiente a supportarlo, perché i risultati di una campagna siano comparabili con quelli delle campagne precedenti bisogna che esista una storia delle attività di marketing memorizzata nel *datawarehouse*.

Entrambe queste due condizioni non sono del tutto verificate per la situazione aziendale in discussione che però ha cercato delle alternative applicabili al suo caso per simulare le metodologie del campione di controllo e di confronto con le passate campagne.

### **Un esempio di analisi di una campagna marketing: il *Trunk Show***

Il *Trunk Show* è una delle campagne marketing più frequentemente ripetute, si tratta di un'attività che consiste nell'invito di un gruppo ristretto di clienti ad una sfilata organizzata ad hoc nei negozi più esclusivi, l'evento avviene due volte all'anno ed ha lo scopo sia di pubblicizzare la collezione sia di fidelizzare la clientela. Il *Trunk Show* è un'attività che viene fatta tradizionalmente nel contesto di business della moda e come molte delle attività di marketing storiche dell'azienda ha alimentato nel tempo l'immagine e il carisma della marca soprattutto in quella cerchia ristretta di clienti riconosciuti come ottimi dall'azienda.

Da quando si è iniziato il processo di automatizzazione del marketing con il progetto di CRM la questione della gestione dei *Trunk Show* è stata guardata da un punto di vista diverso, le domande che ci si è fatti sono:

- *Con quali criteri scelgo chi invitare?*

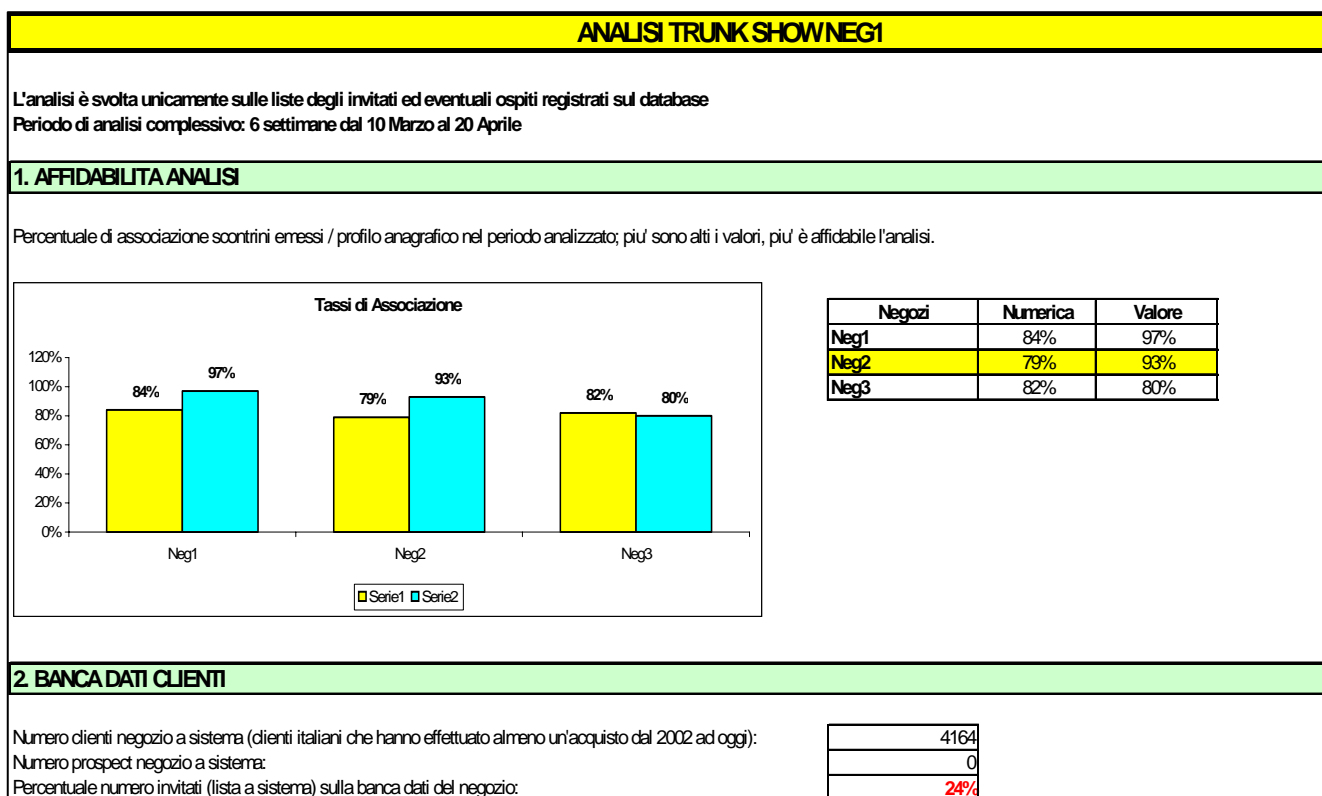
- Come misuro il risultato economico e di immagine della campagna?

Il documento di *reporting* sull'analisi è diviso in 10 sezioni, le prime quattro definiscono le caratteristiche della campagna e ne inquadrano il contesto, le successive quattro sezioni riportano i risultati in termini di fatturato, mentre le ultime due sono di approfondimento. Di seguito verrà esemplificato e discusso il documento di reporting di analisi della campagna.

1. Nella prime due sezioni vengono fatte le considerazioni sull'affidabilità dell'analisi in base alla qualità dei dati. Vengono riportati tassi di associazione cliente-scontrino del negozio su cui si sta conducendo l'analisi e a titolo di confronto i tassi di altri negozi simili.

Nella sezione 2 viene riportata l'ampiezza della banca dati del negozio, il numero cioè, di clienti che hanno effettuato un acquisto registrato in quel negozio.

Queste informazioni servono a dare le informazioni necessarie per valutare statisticamente la rilevanza delle informazioni che emergono dai dati a disposizione.



2. Nella terza sezione vengono descritte le caratteristiche della campagna e dei clienti invitati e viene subito presentata la redemption della campagna.

La *Redemption Contatti* indica il livello di risposta all'invito, quante persone su quelle invitate hanno partecipato al *Trunk Show*. Calcolare questo indicatore non è scontato, come si vede la lista di invito non è unica, esiste infatti una lista di codici cliente memorizzata a sistema, ma gli inviti possono essere anche molti di più, vi si aggiungono spesso anche gli ospiti accompagnatori degli invitati ufficiali ed un certo numero di invitati anonimi non registrati a sistema. Dal momento che la campagna marketing *Trunk Show* ha anche obiettivi di promozione dell'immagine e di fidelizzazione della clientela, il tasso di presenza al *Trunk Show* è un dato che si vuole riportare con la massima correttezza possibile.

Nella sezione viene riportato un altro indicatore di *Redemption* la *Redemption Attività*, che indica il numero di persone che essendo presenti al *Trunk Show* hanno effettuato un acquisto nel negozio in analisi nelle cinque settimane successive all'evento.

Questo indicatore è considerato quello di massima espressività del risultato della campagna in termini di motivazione e stimolo all'acquisto.

Nella sezione 3.a viene riportato un dato di controllo per verificare che i risultati ottenuti siano effettivamente stati influenzati dalla campagna, l'indicatore *%attivati sui non presenti* esprime il numero di clienti attivati nel periodo successivo alla campagna che seppur invitati non hanno partecipato al *Trunk Show*. Questi clienti possono rappresentare una sorta di campione di controllo, perché sono clienti che hanno le stesse caratteristiche di quelli invitati (di fatto sono stati invitati), che non hanno però partecipato alla campagna. Studiare come questi clienti si sono comportati è interessante anche indipendentemente dal fatto di usarli come campione di controllo.

Nella sezione 4 vengono riportate le varie voci di costo per la realizzazione della campagna, la somma delle quali viene divisa per il numero di contatti presenti e non per il numero di contatti totale. In effetti il costo della campagna deve essere visto come il costo di contatto *riuscito* del cliente.

### 3. DESCRIZIONE CAMPAGNA

CODICE CAMPAGNA	TARGET	DATA RICEZIONE INVITO(T)	DATA TRUNK SHOW	lista temporanea da azienda X	presenti non invitati	TOTALE INVITATI	RITORNI POSTALI	CONTATTI EFFETTIVI
00000342	Best Clients	7-mar-05	16-mar-05	30	20	1000	10	990

Inizialmente la lista è stata determinata con un semplice criterio di scelta sul valore storico dell'acquisto dei clienti della boutique. In seguito, la lista è stata modificata dal negozio e si sono pe specificati di selezione. Questa situazione si riflette sicuramente nei risultati (v. punto 5).

La percentuale di ritorno del **mailing** (*redemption lista*) è calcolata dividendo il numero di coloro che hanno presenziato al TS (ospiti esclusi), divisa per il numero di contatti effettivi.

Non avendo dati precisi sulle presenze degli invitati non a sistema, il dato di redemption è ipotetico.

Inoltre, il numero di ospiti presenti deve essere segnalato dal negozio, unitamente ai codici di coloro che è stato possibile registrare a sistema previo consenso.

INVITATI PRESENTI	REDEMPTION CONTATTI	OSPITI REGISTRATI	OSPITI NON REGISTRATI	OSPITI	TOTALE PRESENTI	TOTALE RSVP	TOTALE INVITATI CON 1 OSPITE	TOTALE INVITATI CON PIÙ OSPITI	TOTALE INVITATI SOLI
300	30%	50	10	60	360	0	38	12	41

Gli attivati sono coloro a cui è stata associata una vendita nel brand di riferimento nel periodo analizzato. (Ved elenco nominativi)

La percentuale di ritorno dell'**attività** è calcolata dividendo il totale attivati (sia invitati che ospiti) per il totale presenti.

PRESENTI ATTIVATI	OSPITI ATTIVATI	TOTALE ATTIVATI (CLIENTI + OSPITI)	REDEMPTION ATTIVITA
50	10	60	16%

Si sono inoltre attivate 5 clienti storiche, il cui acquisto non è stato registrato poiché non sono inserite a sistema. Adetta del negozio, 4 di queste clienti avrebbero comunque comprato, indipendentemente dall'invito al TS.

#### 3a Verifica

Nel calcolo di ritorno dell'attività NON vengono presi in considerazione coloro che, pur essendo invitati, non si sono presentati al TS. Alcuni di essi si sono comunque attivati.

Questa verifica funge da campione di controllo.

NON PRESENTI (Ma invitati)	DI CUI ATTIVATI	% ATTIVATI SUI NON PRESENTI
620	35	6%

### 4. COSTO

Il costo contatto è calcolato sulla somma di invitati e ospiti presenti al TS.

SHOW	CATERING	LETTERSHP	AFFRANCATURA	TOTALE	COSTO CONTATTO
6000	€5.000	€1.100	incluso	€12.100	€32



3. La sezione 5 riporta i dati di fatturato generati dalla campagna, generati cioè da quel gruppo di clienti che ha partecipato al *Trunk Show* e ha effettuato almeno un acquisto in quel negozio nelle cinque settimane successive.

Il primo grafico riporta il fatturato totale del negozio nel periodo di analisi e nell'anno precedente nello stesso periodo di analisi.

Successivamente vengono riportati i fatturati settimanali generati dai contatti attivati, e tali fatturati vengono messi a confronto con quelli totali del negozio, l'indicatore si chiama

*%sul Fatturato Neg.*

Viene poi riportato un dato sul comportamento di acquisto dei contatti attivati nel passato, l'anno precedente nello stesso periodo, è un tentativo di approfondire l'esistenza di cause alla base del comportamento non legate all'impatto della campagna.

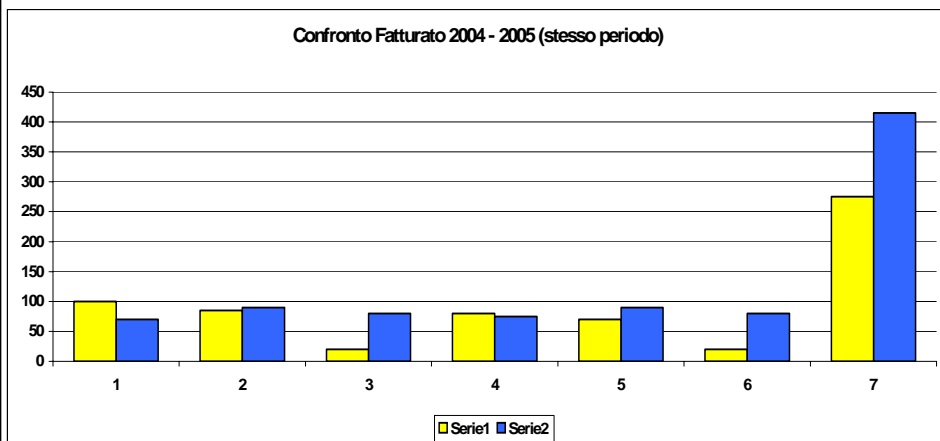
Infine è presentato il fatturato settimanale generato dai contatti non presenti, il nostro campione di controllo.

## 5. FATTURATO

### 5.a Fatturato generale negozio 2005 - 2004

Andamento del fatturato **totale** del negozio Neg1 nella settimana precedente al TS (10-16 Marzo.), nella settimana del TS (17-23 Marzo) e nelle 4 settimane successive (24 Marzo-20 Aprile). Per il 2003 viene preso in considerazione lo stesso periodo calendario.

ANDAMENTO SETTIMANALE						FATTURATO TOTALE	ANNO
SETTIMANA T-1	SETTIMANA T1	SETTIMANA T2	SETTIMANA T3	SETTIMANA T4	SETTIMANA T5		
100	85	20	80	70	€ 20	€ 275	2005
70	90	80	75	90	80	€ 415	2004



### 5.b Fatturato invitati attivati e presenti al TS

Valore totale degli scontrini emessi, al netto di sconti e lordo IVA, per il brand nelle 5 settimane **seguenti** al Trunk Show nel negozio di riferimento Neg1. L'analisi è effettuata sul totale dei **presenti** attivati (invitati più ospiti registrati). A titolo di confronto vengono riportati anche i dati riferiti alla settimana precedente.

ANDAMENTO SETTIMANALE 2005						TOTALE T1 T5	%sul Fatt. Tot. Neg.
SETTIMANA T-1	SETTIMANA T1	SETTIMANA T2	SETTIMANA T3	SETTIMANA T4	SETTIMANA T5	FATTURATO	
€10	€15	€15	€10	€20	€15	€75	27%

### 5.c Numero invitati attivati e presenti al TS

ANDAMENTO SETTIMANALE 2005						TOT T1-5
						ATTIVATI / PRES.
10	5	20	15	10	10	60

### 5.d Spesa media attivati e presenti al TS

ANDAMENTO SETTIMANALE 2005						TOT T1-5
						SPESA MEDIA
€1	€3	€1	€1	€2		€1

Fatturato generato nel 2004 (nel periodo corrispondente all'analisi) dai tutti o alcuni dei contatti attivati per il TS e fatturato generato dagli stessi nel 2005, nel negozio di riferimento. L'affidabilità del dato dipende dai tassi di associazione 2004 (stesso periodo), ivi riportati.

CONTATTI ATTIVATI 2003	FATTURATO 2004	FATTURATO 2005	Negozi	Numerica	Valore
20	€100	€150	NEG1	76%	92%

### 5.e Fatturato invitati attivati non presenti al TS

Fatturato negozio GA Bologna generato da clienti invitati ma non presenti al trunk show (v. 2b)

ANDAMENTO SETTIMANALE						FATTURATO	%sul Fatt. Tot. Neg.
SETTIMANA T-1	SETTIMANA T1	SETTIMANA T2	SETTIMANA T3	SETTIMANA T4	SETTIMANA T5		
10	20	20	10	10	€ 20	€80	29%

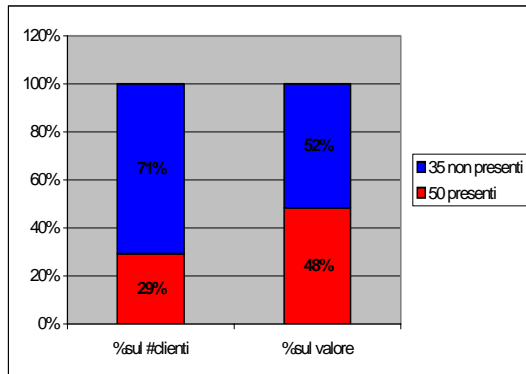
4. Nella sezione 6 viene riportato sotto forma visiva di grafico il paragone tra il fatturato generato e il numero di clienti presenti ed attivati e tra il fatturato generato e il numero dei clienti attivati ma non presenti. Anche questa informazione è di confronto con il nostro campione di controllo, serve per vedere se, in proporzione, gli attivati e presenti (e quindi positivamente colpiti dalla campagna) hanno generato di più rispetto a chi non è stato coinvolto.

La sezione 7 riporta informazioni sugli scontrini medi e le spese medie annuali effettuati dai clienti attivati e presenti, confrontate con lo scontrino medio e la spesa media del negozio.

La sezione 8 rappresenta un approfondimento del comportamento di acquisto e indaga se nello stesso periodo di analisi il cliente ha acquistato anche in negozi diversi del gruppo.

Ed infine segue la lista con i nominativi delle persone che hanno partecipato al Trunk Show ed che si sono attivate.

## 6. CLIENTI MONITORATI: RIPARTIZIONE DEL FATTURATO



## 7. CONFRONTO MEDIE

Scontrino Medio: fatturato generato diviso per il numero di scontrini degli presenti attivati.

Spesa Media: fatturato generato diviso per il numero di presenti attivati.

Scontrino Medio	Scontrino Medio mese precedente	Scontrino Medio Brand Italia	Spesa Media	Spesa Media mese precedente
€100	€80	€90	€200	€150

## 8. COMPORTAMENTO CROSS-BRAND

Numero di contatti attivati ai quali è stata associato un acquisto in uno o più brand diversi dal brand di riferimento e fatturato totale generato.

La somma dei clienti per ciascun brand non è necessariamente eguale al numero totale di contatti attivati.

Brand2	Fatt. Settimana T 1-5
15	€100

## 9. NOTE

Una attività classica di marketing diretto andrebbe valutata anche a confronto con un campione di controllo statisticamente valido. Per un mailing di questo tipo, tale verifica è difficile, dato i numeri esigui del mailing stesso. Detto ciò, in futuro potremmo utilizzare come campione di controllo il comportamento di acquisto dei Best Client di un negozio simile nello stesso periodo.

## 10. ELENCO NOMINATIVI PRESENTI ATTIVI

Codice Cliente	Nome
7000987	
70001046	
70028461	
70082658	
70114015	
70162317	

## 11. ELENCO NOMINATIVI NON PRESENTI, NON REGISTRATI

## Considerazioni sugli indicatori

L'Azienda si sta attrezzando per essere in grado di raccogliere e memorizzare i dati necessari per l'elaborazione di indicatori di *Redemption* delle campagne marketing, tali indicatori infatti sono di grande utilità non solo per l'analisi per la misurazione dei risultati ma anche perchè arricchiscono di molto la possibilità di estrarre conoscenza sul cliente e sull'effetto che le scelte aziendali hanno sul loro comportamento.

Come già detto la tipologia di risposta o *redemption* ad una campagna varia in base alla tipologia di campagna e non è sempre di facile individuazione e definizione, nel caso descritto della campagna tipo *trunk show* una possibile risposta è quella che indica *quante persone si sono presentate al trunk show rispetto a quelle invitate* ma è anche necessario andare più nel dettaglio e chiedersi *quanto hanno generato i clienti invitati al trunk show?*.

Allo stato attuale il programma di gestione campagne permette solo in parte di gestire le risposte, ad esempio nel caso del *trunk show* è possibile memorizzare l'informazione di partecipazione o meno ma per quanto riguarda l'archiviazione di informazioni relativamente al comportamento d'acquisto successivo o contestuale alla campagna non è possibile il monitoraggio né la memorizzazione di dati.

Fino a questo momento gli indicatori sono stati creati di volta in volta nell'ambiente dei documenti di *BusinessObjects*.

L'obiettivo è di giungere alla definizione di un insieme di indicatori standard che possano essere calcolati automaticamente e che possano pertanto essere riutilizzati anche in altre analisi.

Gli aspetti da considerare sono due: la campagna e il cliente rispetto alla campagna.

Per quanto riguarda la campagna marketing dovrebbero essere memorizzati tutti i dati che la descrivono come la tipologia, i clienti coinvolti, il tipo di clienti, i costi, il periodo temporale di riferimento, ecc. e tutti i dati necessari a individuare i risultati ottenuti dalla campagna che variano in base alla sua tipologia.

Per quanto riguarda il cliente gli indicatori che riassumono la reazione alla campagna possono essere considerati come fonte di informazioni molto preziose sul cliente. Gli indicatori sul comportamento verso un certo tipo di campagna potrebbero essere in forma di variabile binaria e indicare, continuando con l'esempio del *Trunk Show*, la partecipazione o l'effettuazione di un acquisto successivamente al *Trunk Show*.

L'utilizzo di tali indicatori potrebbe essere usato per arricchire i dati per la segmentazione della popolazione dei clienti e soprattutto può essere usato per predire le classi di clienti che reagiranno positivamente alla campagna.

#### **4.6 Comportarsi conseguentemente a ciò che si è imparato**

La missione del progetto di CRM afferma di volere ottimizzare la gestione della relazione cliente (finale) nelle aziende del gruppo, in collaborazione con le funzioni interessate, affinché essa si traduca in risultati economici per l'azienda e maggiore soddisfazione per il cliente.

La missione si traduce nella strategia di creazione di una relazione con il cliente sulla base della conoscenza acquisita su di esso grazie alle informazioni reperite durante i contatti tra l'azienda e il cliente, alla condivisione delle informazioni sul cliente di cui l'Azienda dispone e alle attività a lui dirette.

Il passaggio dalla strategia alla definizione degli obiettivi è il frutto di una serie di contrattazioni che dipendono dal contesto di partenza in cui si è inserito il progetto.

Il progetto è partito con il primo obiettivo di dimostrare che ci fosse effettivamente in Azienda la necessità di adottare un approccio di tipo CRM e che il contesto organizzativo fosse in grado di adattarsi ai cambiamenti aziendali necessari per l'applicazione della strategia.

Nonostante la cautela nei confronti del progetto il management d'Azienda ha dimostrato una certa sensibilità verso le caratteristiche del CRM e in particolar modo verso il fatto che il CRM serve per dare delle informazioni ed indicazioni di tipo quantitativo al livello direzionale.

La volontà di approcciare almeno alcuni aspetti del business in modo *quantitativo* è una condizione necessaria perché sia possibile che il ciclo di monitoraggio e analisi si chiuda con l'azione.

Buona parte delle attività necessarie per il raggiungimento degli obiettivi che riguardano la preparazione del sistema informatico per il monitoraggio, la memorizzazione e l'organizzazione dei dati non produce risultati di interesse per il top management ma riguarda i manager di area.

Ad esempio l'automatizzazione di una funzione di cassa può essere molto utile per il cassiere, essere rilevante per il direttore del negozio e, forse, per il manager *Retail* ma non produrrà nessun tipo di risultato apprezzabile a livello strategico.

I risultati delle analisi sui dati invece, innanzitutto testimoniano nel complesso la capacità dell'intero sistema di produrre informazioni di tipo direzionale e poi producono risultati di utilità per il top management.

Dal momento che le analisi condotte sono di tipo OLAP, i risultati forniscono delle descrizioni basate sui dati, quindi descrizioni di tipo quantitativo del business.

Questo tipo di informazioni servono per confermare o meno una conoscenza supportata precedentemente dall'esperienza ma non dai dati, per verificare un'intuizione, o per misurare quantitativamente un fatto.

In sintesi la conoscenza scaturita dall'analisi dei dati supporta i manager nel prendere decisioni fornendo una misurazione quantitativa di alcuni fatti del business.

Vista l'importanza delle decisioni prese sulla base delle analisi sul datawarehouse la qualità dei dati è determinante; i dati devono fornire una rilevanza statisticamente significativa, devono essere cioè sia nel numero sia nel livello di correttezza in grado di rappresentare la realtà senza distorsioni.

La qualità del dato riguarda gli obiettivi operativi del progetto, nelle fasi di raccolta dei dati e integrazione delle fonti, le attività a supporto di queste fasi sono di training e incentivazione degli utenti (negozi e Press/Promozioni) e a livello di negozio si è tradotto nella decisione di inserire tra gli obiettivi di performance anche quello di abbinamento scontrini.

L'obiettivo di qualità dei dati pur essendo di tipo operativo coinvolge il processo in tutte le sue fasi ed attività ed ha dato luogo a riflessioni rilevanti anche a livello strategico, uno degli argomenti più discussi è l'opportunità o meno di inserire un meccanismo tipo carta punti o carta fidelity. Il meccanismo supportato da una carta magnetica che automaticamente identifica il cliente al momento dell'acquisto risolverebbe buona parte dei problemi di qualità dei dati che l'Azienda incontra.

L'introduzione di una carta fidelity oltre a presentare degli aspetti operativi non trascurabili come la necessità di un sistema di riconoscimento magnetico alle casse e la modifica del programma di gestione delle attività delle casse, coinvolge anche aspetti

strategici. Le riflessioni riguardano il fatto che un tale meccanismo sia adeguato all'immagine dell'azienda e che venga accettato positivamente dal cliente.

Per rispondere a questi dubbi, soprattutto a quelli legati alla possibile reazione del cliente, il team di progetto CRM, ha ideato una campagna marketing che introducesse in un arco di tempo limitato la presenza della carta punti. Si è trattato di un vero e proprio test, le domande a cui si è cercato di dare delle risposte in base ai risultati della campagna messa in atto sono del tipo

- Target clienti: a quale tipologia di cliente proporre una carta fidelity ?
- Tipologia di negozio: quale tipologia di negozio e quindi quale tipologia di cliente è più adatto?
- Che cosa offrire e con quali modalità: quale meccanismo proporre, carta di accumulo punti in base agli acquisti, carta fidelity di altro genere, ecc.?

L'esempio riportato esemplifica come una esigenza scaturita o comunque messa in evidenza dal progetto di CRM possa diventare un'esigenza percepita dall'azienda, nei vari livelli che la compongono, e come il ciclo di monitoraggio, analisi e azione si sviluppi in maniera circolare e possa supportare il processo decisionale.

Come esemplificato nella piramide delle fasi critiche del CRM in figura 9, la punta della piramide indica la volontà da parte del manager di condurre un'indagine di soddisfazione del cliente a coronamento e completamento delle attività di business intelligence sui dati raccolti.

La soddisfazione del cliente appare esplicitamente nella definizione della missione del progetto di CRM.

La decisione di condurre un'indagine di soddisfazione del cliente è nata anche dai risultati delle analisi OLAP sul comportamento dei clienti che sono state mostrate al *management*, le analisi infatti consistono nella descrizione di alcuni aspetti del business ma non ne danno spiegazioni dei motivi e delle cause, fanno emergere aspetti critici o di opportunità ma è l'azienda che poi deve decidere che cosa fare di queste informazioni.

L'indagine di soddisfazione del cliente voluta dall'Azienda prende in considerazione sia gli aspetti quantitativi del comportamento del cliente che emergono dall'elaborazione



dei dati del *datawarehouse* sia gli aspetti quantitativi che emergono durante l'indagine diretta ai clienti sia quelli qualitativi che emergono nell'interazione.

L'indagine di soddisfazione cliente è in linea con la strategia di CRM perché implica il contatto diretto con il cliente, e la memorizzazione ed elaborazione dei dati quantificabili che riguardano l'interazione ma anche l'elaborazione delle informazioni qualitative.

L'indagine apporta conoscenza durante tutte le sue fasi, le informazioni sul cliente che l'azienda sta cercando.

La presenza dell'imprenditore nel processo decisionale può dare una direzione diversa da quella che verrebbe presa se fossero solo i manager a decidere, l'imprenditore infatti non applica sempre il criterio economico.

Nell'Azienda in questione le qualità, il grande intuito e carisma dell'imprenditore guidano l'azienda nelle sue scelte e può accadere che alla fine del processo vengano prese delle decisioni in contraddizione con i dati emersi dall'analisi.



## **5. Idee per estendere l'ambienti di Business Intelligence**

### **5.1 L'ambiente di business intelligence attuale**

Nel capitolo precedente è stata fatta la descrizione dell'esperienza di avvio della costruzione di un ambiente di business intelligence per il CRM di un'azienda che opera nel settore moda.

E' stata ricostruita la vita del progetto di CRM ripercorrendo cronologicamente le fasi critiche e le attività di ciascuna fase, riflettendo sui problemi che si sono presentati di volta in volta e sulle soluzioni scelte e analizzando il ciclo di monitoraggio ed elaborazione dei dati che si conclude e si riapre, vista la sua natura ciclica, con la produzione di informazione direzionale e con la conseguente azione da parte del management.

L'esperienza descritta rappresenta un caso di primo approccio alla business intelligence, il progetto pilota ha avuto l'obiettivo di dimostrare che:

1. l'Azienda ha bisogno di un tale sistema;
2. l'Azienda è in grado di implementare tale sistema ;
3. se i primi due punti sono verificati, preparare il terreno per un ambiente di business intelligence completo.

Il punto 1 è stato ampiamente discusso nel capitolo precedente, i risultati portati dal progetto pilota hanno giustificato ed incoraggiato la prosecuzione del progetto sotto condizioni diverse, cioè come progetto sostenuto dal consenso del management e con un budget superiore a quello precedentemente deciso.

Per quanto riguarda il punto 2, il progetto pilota ha dimostrato che l'Azienda ha le capacità tecniche e gestionali e le risorse per implementare e rendere attuabili le seguenti attività:

- il monitoraggio semi-automatizzato delle interazioni con il cliente in negozio;
- la raccolta dei dati sul cliente nella base di dati operativa;

- l'organizzazione di tali dati secondo una logica multidimensionale;
- l'analisi dei dati con strumenti di analisi tipo OLAP
- la produzione di documenti di *reporting*.

Alcune di queste attività sono da perfezionare e da completare come è emerso nella descrizione del capitolo 4, come ad esempio il fatto che il datawarehouse non è stato sviluppato in un ambiente dedicato o come il fatto che il monitoraggio delle interazioni con il cliente in realtà si riferisce alle sole interazioni che avvengono in negozio e non è ancora possibile monitorare i clienti anche attraverso altri canali.

Il sistema però ha dimostrato di servire al suo scopo, cioè di fornire dei dati corretti (rispetto alla sorgente) raggruppabili secondo la logica cliente-centrica e analizzabili; i risultati delle analisi su tali dati ha prodotto informazioni utili e di supporto per la presa di decisioni.

Inoltre, nella logica di sviluppo incrementale le modifiche, i miglioramenti e le aggiunte sono da considerarsi parte del processo di sviluppo.

Il sistema descritto nel capitolo 4 pone delle buone basi per la costruzione di un ambiente di *business intelligence*, per il completamento dell'ambiente, su queste basi deve poggiare un sistema di *data mining*.

Nel progetto pilota l'analisi dei dati si è limitata all'analisi di tipo statistico descrittivo dei fatti secondo la logica multidimensionale e con l'uso di strumenti di *query OLAP*.

Il data mining è qualcosa di molto più complesso di una tecnica di analisi e, al di là degli slogan pubblicitari, è difficile da implementare e difficile da gestire ma può portare a risultati molto importanti con un impatto (positivo) sul business molto forti.

*Il data mining consiste nel processo di esplorazione ed analisi, in maniera automatica o semi-automatica, di grandi quantità di dati al fine di scoprire regole e percorsi interessanti nei dati.*<sup>4</sup> Il data mining è un processo che estrae conoscenza

---

<sup>4</sup> Si veda M.J.A Berry e G.S. Linoff "Data Mining Techniques For Marketing, Sales, and Customer Relationship Management" Second Edition, Wiley Publishing, Inc.

attraverso l'individuazione di regole e associazioni tra i dati non note e non visibili ad occhio nudo.

La differenza tra i risultati delle analisi OLAP e il data mining è che le analisi OLAP permettono di produrre rapporti di sintesi a partire da interrogazioni precise, mentre i sistemi di data mining forniscono una serie di modelli che evidenziano aspetti di qualche interesse nei dati e che erano precedentemente sconosciuti.

Un sistema OLAP dà una risposta a domande tipo: *trovare i clienti che si sono attivati a seguito di una partecipazione al trunk show e che hanno speso più di 100 euro in accessori.*

Un sistema di data mining invece dà risposta tipo: *trovare il profilo generale dei partecipanti al trunk show che si sono attivati.*

Per capire se e perché l'Azienda in questione dovrebbe avere bisogno di un sistema di data mining è necessario capire che tipo di risposte può dare un sistema di data mining.

Il data mining applicato nel contesto del marketing è usato per migliorare il business sia dal punto di vista del costo sia dal punto di vista della redditività; il data mining è infatti usato per diminuire i costi delle campagne marketing evitando di contattare clienti che probabilmente risponderanno negativamente ma è anche usato per incrementare la redditività dei clienti attraverso la calibratura dell'offerta su quelle classi di clienti con certe caratteristiche che indicano che sono pronti per essere spinti verso un comportamento più profittevole (*cross brand o up selling*).

Il CRM tende, idealmente, alla gestione uno-a-uno del cliente; un buon sistema di gestione della relazione con il cliente ha bisogno di capire chi sono i propri clienti e che cosa vogliono o non vogliono, ciò significa anticipare i loro bisogni e andare loro incontro in maniera proattiva.

Il data mining quindi si adatta in maniera perfetta a supportare un approccio di CRM.

Una volta capito se un sistema di data mining possa fare al caso del proprio business bisogna capire e valutare se l'Azienda dispone delle risorse e delle capacità necessarie ad implementarlo e se il costo e i rischi di adozione di un tale sistema siano pienamente ripagati dai benefici che genererebbe.

## 5.2 L'ambiente di business intelligence verso cui si vuole andare

Come discusso nel capitolo 2 l'impresa cliente centrica considera il cliente come la fonte primaria di valore, il valore che il cliente ha per l'azienda però cambia a seconda della fase della relazione in cui si trova e questo fatto è rappresentato dal ciclo di gestione del cliente.

Esiste un altro ciclo altrettanto importante ed è il ciclo di vita della persona, il ciclo di vita di ogni persona è unico e scandito da fasi, tappe e traguardi che possono essere molto diversi da persona a persona, in generale, ci sono molti eventi importanti nella vita di ognuno che condizionano anche le scelte economiche e di consumo.

Il CRM combina la consapevolezza dei cicli di vita delle persone con il ciclo di vita della relazione tra il cliente e l'azienda.

La business intelligence supporta l'azienda cliente centrica che cerca da un lato di conoscere i consumatori e riconoscere le loro caratteristiche e tipologie e dall'altro cerca sulla base della conoscenza della propria clientela di fare l'offerta giusta alla persona giusta, nel momento giusto e al prezzo giusto.

Affrontiamo, per il momento, in maniera separata la gestione dei due cicli, per primo vediamo il ciclo di gestione del cliente, in questo contesto il cliente è visto come un'entità che risponde alle azioni dell'impresa. Questo ciclo della vita del cliente è molto importante perché è direttamente legato al valore che il cliente ha per l'azienda.

Il ciclo di vita del cliente si svolge nelle fasi di:

- *Prospect*: persona che non è ancora cliente ma che appartiene al segmento di mercato in cui opera l'azienda.
- *Responder*: persona che non ha ancora effettuato il suo primo acquisto ma ha dimostrato un interesse specifico per l'Azienda.
- *Cliente*: persona che ha consumato il prodotto o il servizio.
- *Ex clienti*: cliente che ha abbandonato gli acquisti.

Nel caso dell’Azienda oggetto di studio le fasi del ciclo di vita potrebbero essere le seguenti: i *prospect*, cioè quelle persone che non sono ancora clienti ma che hanno le caratteristiche per diventarlo, sono ad esempio i clienti delle imprese concorrenti; i *responder* sono persone che sono entrate in negozio ed hanno firmato il modulo privacy pur non avendo effettuato nessun acquisto. I clienti sono coloro che hanno rilasciato il consenso al trattamento dei loro dati personali e hanno effettuato acquisti recentemente, infine gli ex clienti sono quelle persone che erano clienti ma che da un certo momento in poi non hanno più effettuato acquisti.

Gli eventi che segnano il ciclo di vita del cliente sono eventi provocati dall’azienda, l’azienda mira le sue azioni in maniera diversa sui diversi gruppi di persone (*prospect*, *responders*, clienti, ex clienti). In queste fasi l’offerta si dimostra tanto più efficace quanto più accuratamente è mirata sul target giusto di persone e clienti; la business intelligence e in particolare il data mining supportano e forniscono le informazioni per far questo.

- Azioni di acquisizione: i destinatari sono i *prospect*.

Le attività possono consistere in campagne marketing mirate su certi target di clienti o campagne pubblicitarie destinate al pubblico di massa attraverso i media.

Nel caso in cui la campagna venga circoscritta a un determinato target la business intelligence può essere di grande aiuto nell’individuare le caratteristiche che fanno dei *prospect* dei futuri clienti, ad esempio attraverso l’analisi dei profili assunti dai *responders* o dai clienti prima di avere risposto o acquistato.

- Azioni di attivazione: i destinatari sono i *responder*.

Quando un *prospect* diventa un *responder*, attraverso la sottoscrizione di un modulo ricevuto per posta, sul web o trovato in negozio, la persona perde la sua anonimità e diventa un’entità vera e propria per l’azienda. Si può presumere che i *responder* siano le persone che più probabilmente diventeranno dei clienti, lo studio e l’analisi delle caratteristiche di queste persone è molto utile per costruire dei modelli predittivi su quali *prospect*

diventeranno *responder* ed inoltre il monitoraggio dei *responder* che effettivamente si attiveranno sarà indispensabile per costruire i modelli predittivi per determinare le caratteristiche dei *responder* che diventeranno clienti.

- Azioni di gestione della relazione (ottimizzazione della relazione): i destinatari sono i clienti.

I clienti attivi sono la fonte principale di informazione, le informazioni sono prodotte dalla memorizzazione delle transazioni di acquisto e dalla memorizzazione delle risposte alle attività di marketing proposte dall'azienda.

Le attività possono essere di incoraggiamento all'acquisto o di incoraggiamento al comportamento *cross selling* e di spinta verso offerte superiori (*up selling*). In questa fase il cliente si può trovare in stati mentali ed emotivi diversi rispetto alla relazione con l'azienda (soddisfazione, fedeltà mentale, ecc.) e può trovarsi in stati diversi della sua vita.

L'analisi del comportamento di acquisto può portare all'individuazione di gruppi o segmenti di clienti diversi. Ad esempio a seconda della tipologia dei primi acquisti, caratterizzati dal numero di prodotti, tipo di prodotti, ammontare speso, frequenza di acquisto ecc., è possibile predire il comportamento futuro.

Anche i dati relativi ai comportamenti di acquisto dei clienti persi sono molto importanti in questa fase perché identificando e isolando le caratteristiche di quei clienti che hanno abbandonato l'azienda si può predire tipologie di comportamenti da considerarsi a rischio di abbandono.

Il secondo ciclo di vita del cliente che qui si vuole descrivere è quello che riguarda l'individuo, la vita delle persone infatti è scandita da eventi che cambiano il modo di pensare e di agire delle persone nel tempo e che pertanto condizionano anche i loro stili di vita e di consumo. Eventi come l'inizio dell'università, il primo lavoro, il matrimonio e la pensione cambiano radicalmente e piuttosto repentinamente le abitudini delle persone e un'impresa cliente-centrica deve essere in grado di saper gestire il ciclo di vita



di ogni cliente, di anticiparlo e di sfruttare i generatori di valore che in ogni stadio del ciclo si presentano a suo favore.

Il valore del cliente dipende fortemente dal suo stato interno che caratterizza ogni momento della relazione con l'azienda; questo fatto porta l'impresa a dover allargare il suo campo di osservazione ed analisi anche a ciò che succede nella vita di ogni individuo.

Presa coscienza da parte dell'impresa del fatto che il cliente non è solo un'entità che risponde alle sollecitazioni del mercato, la relazione che questa deve cercare di creare deve avere lo scopo di capire il cliente, anche come individuo, attraverso l'instaurazione di un processo dinamico e interattivo.

Il valore del cliente rapportato alla catena della redditività descritta nel secondo capitolo rientra in gioco, la relazione tra il cliente e l'impresa può trovarsi in una delle cinque fasi di soddisfazione, fiducia, fedeltà comportamentale, fedeltà mentale e lealtà.

Nella trattazione sopra sono stati evidenziate le relazioni tra CRM e data mining, sintetizzando, il data mining può essere applicato in ogni fase del ciclo di vita del cliente e del ciclo di vita della relazione tra cliente ed impresa per approfondire la conoscenza del cliente e per migliorare l'offerta dell'azienda.

Le opportunità di utilizzo del data mining sono molteplici e le attività descritte fin qui si possono riassumere nella segmentazione e nella previsione delle risposte del cliente alle attività proposte dall'impresa.

Ritornando allo studio del nostro caso pratico per il momento sono stati tracciati, in maniera generica, gli obiettivi di business ma rimane ancora da analizzare in profondità il processo di data mining da mettere in atto. Dall'analisi del processo di data mining dovranno emergere con chiarezza:

1. gli obiettivi di business;
2. la definizione dei relativi problemi di data mining;
3. le risorse (tecnologiche, umane, finanziarie) disponibili e non;
4. lo stato dei dati.

Nel paragrafo seguente viene analizzato ognuno di questi quattro punti in riferimento alla situazione dell'Azienda del nostro caso di studio.

### **5.3 Analisi dei requisiti delle nuove funzionalità dell'ambiente di *business intelligence***

L'Azienda del nostro caso di studio ha deciso di dotarsi di un ambiente vero e proprio di business intelligence a supporto della strategia di CRM; tale ambiente è prevalentemente costituito dalle tecnologie di *datawarehousing* e di data mining.

In Azienda esiste già un *datawarehouse* del cliente, la tecnologia di datawarehouse è conosciuta e le analisi di tipo statistico descrittivo sul datawarehouse sono ormai un'attività assodata mentre il data mining è ancora una tecnologia del tutto sconosciuta.

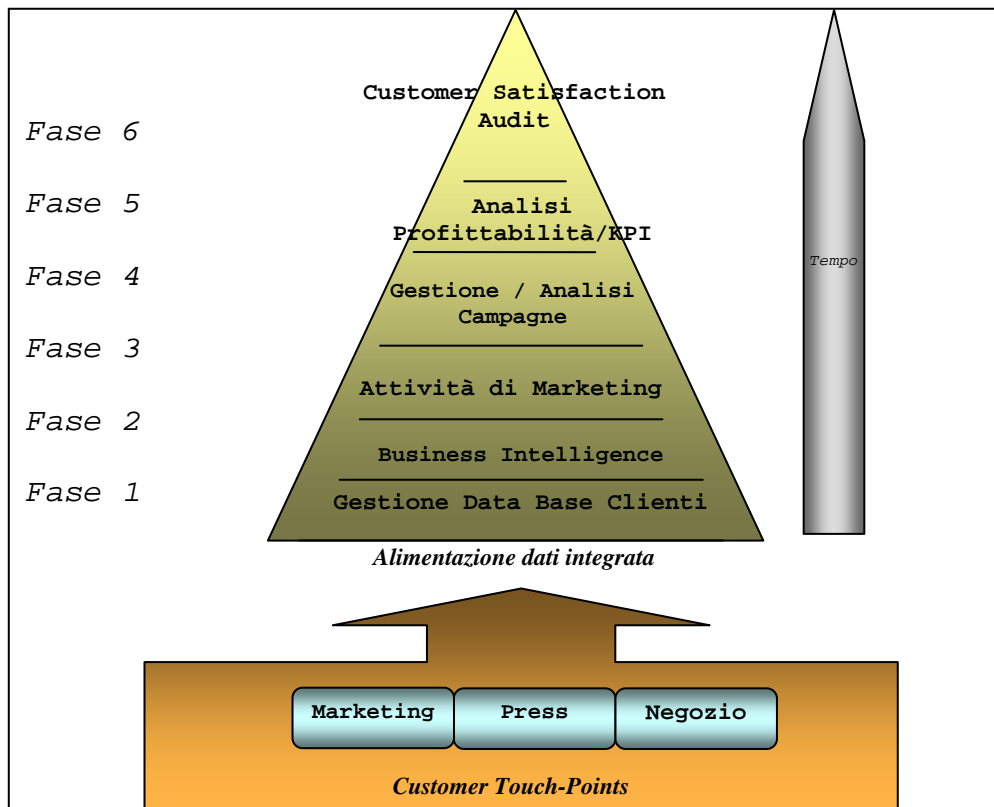
I motivi per cui l'Azienda vuole completare l'ambiente di business intelligence con l'introduzione del data mining sono legati alla volontà di muoversi verso una realtà cliente-centrica basata sul valore del cliente e alla consapevolezza che l'acquisizione della conoscenza del cliente passa attraverso l'utilizzo di tutte le tecnologie di *business intelligence*. Le considerazioni fatte nei paragrafi precedenti di questo capitolo valgono anche per l'Azienda. L'aspettativa più forte che l'Azienda ha nei confronti del data mining è che esso dia la capacità di approfondire la conoscenza della propria clientela indagando sia gli aspetti individuali della vita dei clienti sia le tipologie di rapporto che l'azienda ha instaurato con loro.

Il data mining è un processo costoso e delicato e qualsiasi azienda che si voglia avvicinare al processo deve innanzitutto chiarire e definire che cosa vuole fare con il data mining e successivamente, in base agli obiettivi di business definire le fasi del processo di implementazione.

#### **Gli obiettivi di business**

Ritorniamo sulla piramide delle fasi critiche del CRM tracciata dal manager di progetto, alla base della piramide appena sopra la gestione del database clienti si trova il livello di

*business intelligence*, che supporta i livelli superiori di: *attività di marketing*, *gestione/analisi campagne marketing*, *analisi profittabilità/KPI* ed, infine, *Customer Satisfaction Audit*.



**Figura 15: il progetto di CRM per fasi critiche**

Gli obiettivi di business, in generale, sono:

- l'ottimizzazione delle campagne marketing;
- l'ottimizzazione del valore del cliente.

I due obiettivi non sono disgiunti ma fortemente interdipendenti e sono obiettivi di ampio raggio, coinvolgono cioè l'azienda nel suo complesso e necessitano di un processo continuo.

L'ottimizzazione delle campagne coinvolge innanzitutto il concetto di costo e di budget, dal momento che ogni campagna marketing ha un budget limitato (solitamente esprimibile come numero contatti x costo contatto) ha senso cercare di focalizzare la propria attività su quei clienti che con buona probabilità daranno una risposta positiva.

Ciò significa, in termini di data mining, che si possono mettere in atto due processi: costruire un modello predittivo, basato sui dati dei comportamenti di acquisto nel passato, per determinare chi risponderà positivamente all'offerta e/o, attraverso l'esplorazione dei dati e tecniche di data mining non diretto, individuare quei gruppi di clienti che risponderanno con buona probabilità all'offerta.

L'ottimizzazione del valore del cliente riguarda ancora le campagne marketing ma in senso più ampio, riguarda cioè la gestione complessiva di tutte le attività di marketing che deve essere finalizzata alla massimizzazione del valore di ogni cliente. In termini di data mining significa partire dalla segmentazione dei clienti, considerando congiuntamente il cliente e le campagne marketing; la segmentazione della popolazione dei clienti serve per limitare gli spazi di ricerca delle soluzioni. La segmentazione sarà fatta sui dati che riguardano il cliente come individuo e sui dati che lo riguardano come consumatore e partner nella relazione con l'Azienda.

In che senso la segmentazione serve per limitare lo spazio di ricerca? Il metodo individua gruppi di clienti simili tra di loro e dissimili con i clienti degli altri gruppi, ogni gruppo di clienti è caratterizzato da una serie di comportamenti e caratteristiche socio-demografiche simili, è possibile trattare ogni gruppo come un cliente diverso, prescindendo dal fatto che in realtà il gruppo è un insieme di clienti. Le attività di marketing saranno centrate sui segmenti di clienti a seconda di quali sono le caratteristiche del gruppo e delle loro risposte predette.

La segmentazione serve anche sia come fonte di conoscenza sul cliente sia per trovare una definizione da condividere su chi è il cliente per l'Azienda e qual è il suo valore. L'azienda già con l'applicazione del metodo RFM aveva iniziato un'attività di *scoring* con l'obiettivo di trovare un criterio per assegnare un valore ad ogni suo cliente.

Inoltre il fatto che negli obiettivi di business sia prevista come punta della piramide l'indagine di soddisfazione del cliente denota che l'Azienda sta cercando di individuare i driver della soddisfazione e degli altri stati del cliente per capire in quale fase della relazione ogni cliente si trova, soddisfazione, fiducia, fedeltà mentale, fedeltà comportamentale e lealtà e per gestire le fasi della relazione in maniera appropriata.

## **Gli obiettivi di data mining**

Dagli obiettivi di business discussi sopra emerge che l'ambiente di *business intelligence* in grado di supportare il business deve comprendere le funzionalità di:

- segmentazione implementata con tecniche di apprendimento non supervisionato (*clustering*, regole associative) per l'individuazione di gruppi di clienti con caratteristiche simili sia dal punto di vista del comportamento d'acquisto sia dal punto di vista delle caratteristiche individuali;
- segmentazione implementata con tecniche di apprendimento supervisionato (ad esempio alberi di decisione) per descrizione dei cluster ed approfondimento della conoscenza delle loro caratteristiche;
- previsione risposte dei cluster alle campagne con tecniche di apprendimento supervisionato (alberi di decisione), la previsione avviene sulla base dello studio dei comportamenti passati dei clienti rispetto alle campagne a cui hanno partecipato.
- segmentazione dei clienti per stadi della relazione sulla base dei driver della soddisfazione individuati nella *Customer Satisfaction Audit* con tecniche di apprendimento supervisionato (alberi di decisione).

## **Le risorse**

Le risorse necessarie per costruire l'ambiente di *business intelligence* appena definito in parte sono già presenti in Azienda e in parte devono essere acquistate o sviluppate internamente, sono di vario genere, risorse umane, organizzative, tecnologiche e finanziarie.

Le competenze organizzative per la gestione dello sviluppo dell'ambiente di *business intelligence* che l'Azienda si è riproposta di fare sono detenute da vari attori aziendali, innanzitutto il progetto è sponsorizzato da un top manager che è colui che in ultima analisi dedica le risorse e risponde dei risultati del progetto nella sua globalità.

All'interno del team progetto di CRM, il manager di progetto ha il compito di interpretare le aspettative della direzione nei riguardi dell'ambiente e di individuare i relativi obiettivi di business, il manager avrà il compito di organizzare e gestire il processo e quindi di preparare una proposta di processo di data mining dove indicherà alcune tra le alternative di soluzioni possibili.

Per arrivare all'elaborazione di una proposta il manager di progetto dispone delle conoscenze di *business intelligence* dell'analista dei dati e può ricorrere a pareri e consulenze esterni e dispone dei soliti esperti di datawarehouse e dell'infrastruttura tecnica.

Le aree coinvolte dall'ampliamento dell'ambiente restano le stesse della prima fase del progetto, la *Retail*, il marketing, la *Press/Promozioni* e i sistemi informativi/organizzazione; in ogni area i ruoli con capacità decisionale, per quanto riguarda il progetto di CRM, rimangono gli stessi e affidati alle stesse persone.

Nel caso si volesse introdurre nel progetto di CRM la gestione di altri canali di contatto con il cliente, come il *call center* o il sito web potrebbe essere necessario coinvolgere nuove persone nel progetto.

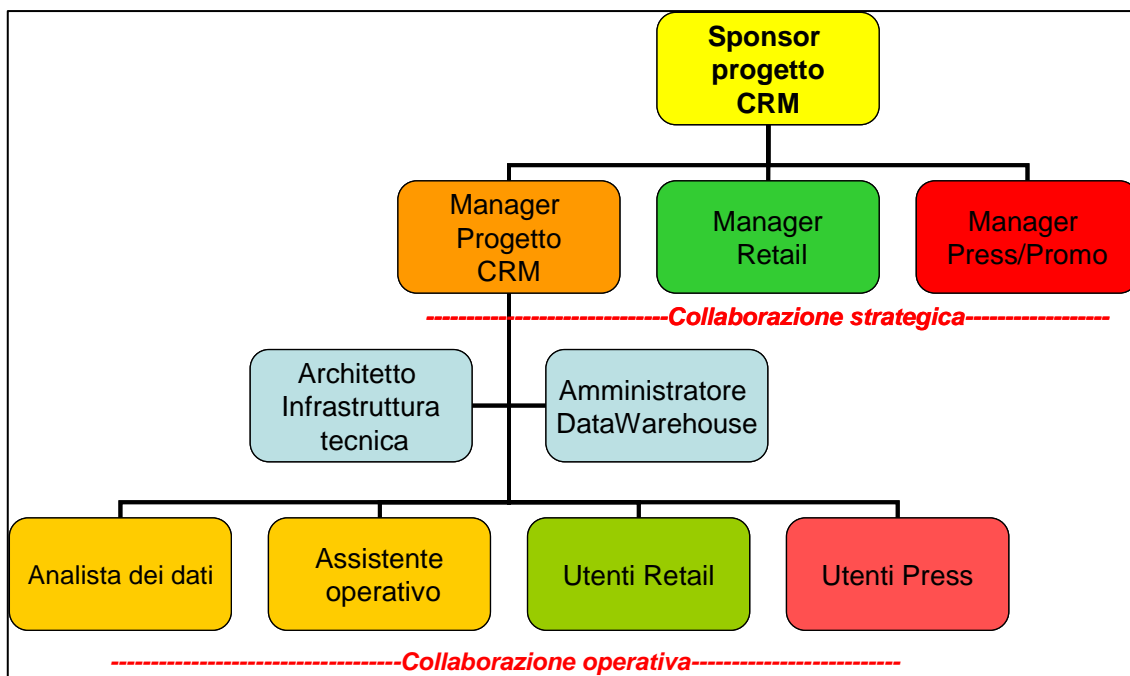


Figura 16: organigramma del progetto CRM

Per quanto riguarda le risorse di tipo tecnologico l'Azienda dispone già di una base di dati operativa dove sono gestiti i dati che riguardano i clienti e le campagne marketing, dispone quindi anche delle persone con le competenze necessarie sia per la manutenzione sia per un eventuale sviluppo. Dispone delle persone con competenze specialistiche di *datawarehousing* che sono le stesse persone che hanno messo a punto l'ambiente intermedio di datawarehouse e l'ambiente di analisi con *BusinessObjects*. L'Azienda ha già preventivato la migrazione dei dati su un ambiente dedicato di *datawarehouse* e manterrà sul nuovo ambiente lo strumento di analisi OLAP e reporting *BusinessObjects*.

Per quanto riguarda l'introduzione di un sistema di data mining con gli obiettivi definiti in precedenza sono necessari però degli adeguamenti tecnologici, che possono riguardare il programma di gestione delle casse, la base di dati, il *datawarehouse* e l'introduzione di nuovi software. Gli adeguamenti del sistema informatico già esistente non dovrebbero però presentare dei grossi problemi anche se potrebbero richiedere del tempo, tutti gli adeguamenti sono comunque provocati dalla necessità di reperire nuovi dati o trasformare quelli esistenti in maniera tale che l'indagine sia possibile non solo sui dati del venduto in negozio ma anche su altri dati che riguardano il cliente. Nella sezione successiva si discuterà a fondo dell'aspetto dei dati nel progetto.

Ciò che invece è necessario sviluppare in Azienda sono gli strumenti e le competenze necessarie per utilizzare al meglio la tecnologia di data mining.

Quando ci si incammina per una nuova strada, che inizialmente non rientra nelle competenze *core* del business dell'impresa, una delle prime riflessioni riguarda l'opportunità di sviluppare internamente il sistema che manca o delegarne lo sviluppo all'esterno.

La decisione *make or buy* è solitamente frutto di una riflessione profonda all'interno dell'organizzazione che prende in considerazione aspetti economici, aspetti di fattibilità tecnica (come quelli di cui stiamo parlando in questo paragrafo) e aspetti strategici e di identità dell'azienda.

Nel caso del data mining la decisione *make* equivale al reperimento di uno strumento software di modellazione e alle capacità e conoscenze necessarie per usare tale strumento e gestire l'intero processo di data mining.

La decisione *buy* equivale invece ad acquistare un servizio di data mining che esclude l'acquisto dello strumento e delle competenze di data mining, tale servizio dovrà però essere condiviso e replicabile nel tempo.

I vantaggi della prima soluzione sono che viene introdotto in azienda la tecnologia del data mining e viene sviluppata una conoscenza interna degli strumenti e della teoria ed inoltre viene promossa una sensibilità dell'organizzazione verso i concetti alla base di tale tecnologia. L'azienda inoltre mantiene il controllo sull'intero processo di generazione della conoscenza assicurando una coerenza di fondo di gestione del processo. Gli svantaggi di una tale soluzione sono gli oneri connessi all'acquisizione degli strumenti e allo sviluppo delle competenze e i rischi: rischio di inadeguatezza della struttura informatica esistente, costi di un eventuale dismissione del sistema, ecc.

La soluzione *buy* presenta i vantaggi di non impegnare in alcun modo l'organizzazione nel processo se non nella cessione dei dati e nell'acquisizione dei risultati, gli svantaggi riguardano però il fatto di non controllare e pilotare in prima persona il processo, non sviluppare una conoscenza della tecnologia all'interno, di essere estremamente dipendenti dal fornitore del servizio con tutti i rischi annessi. Nei casi in cui il progetto di data mining sia centrato su un aspetto specifico e estemporaneo del business che si vuole analizzare una sola volta, la soluzione di delegare totalmente il processo all'esterno è probabilmente la migliore che si possa prendere.

Esistono tutta una serie di soluzioni ibride tra lo sviluppo interno e l'outsourcing totale che possono essere applicate nel caso del data mining e che riguardano essenzialmente la possibilità di dotarsi da parte dell'Azienda della tecnologia e di ricorrere a competenze esterne che gestiscono il processo di data mining a vari livelli. In questi casi quindi devono essere prese due decisioni, una che riguarda che tipo di strumento acquistare e l'altra che riguarda la tipologia di competenze esterne di cui si vuole usufruire.

Riguardo allo strumento di data mining occorre innanzitutto chiarire che la decisione di quale strumento acquistare è importante ma che si tratta solo di una piccola parte del sistema di *business intelligence*, acquistare uno strumento di data mining senza considerare il processo nella sua interezza non risolve il problema del data mining.



La differenza principale tra i software di data mining in commercio sta nel fatto che alcuni si basano su modelli predefiniti ed altri permettono invece la costruzione del modello di data mining che l'azienda vuole implementare.

I programmi che offrono dei modelli statici possono essere basati su problemi di business specifici di un certo contesto di business o settore oppure possono fornire una serie di strumenti applicabili ai problemi di business più diffusi, indipendentemente da un contesto specifico. I vantaggi di una soluzione basata su modelli è che vengono automatizzata e standardizzata la maggior parte delle fasi del processo di data mining. Gli svantaggi sono dovuti essenzialmente alla rigidità dei modelli che dovendo avere un carattere di generalità non possono necessariamente essere del tutto conformi ad ogni situazione aziendale in cui vengono applicati.

I programmi che invece offrono strumenti per la costruzione di modelli offrono la massima libertà all'analista sulla scelta di ogni componente in ogni fase del processo, alcuni strumenti oltre ad avere un'interfaccia grafica danno anche la possibilità di modificare e creare delle nuove funzionalità scrivendo codice con i linguaggi di programmazione più conosciuti o con uno pseudo codice del programma.

Lo svantaggio principale della scelta di uno strumento di questo genere è che è impossibile automatizzare la maggior parte del processo di data mining, come le fasi di individuazione degli obiettivi di business, individuazione dei relativi problemi di data mining, la preparazione dei dati, la modellazione e l'interpretazione dei risultati. Oltre ad un fatto di efficienza il processo di data mining è lungo, costoso e molto delicato, alla fine del quale si può anche arrivare a dei risultati inconsistenti, sbagliati o di nessuna utilità; questo modo di lavorare per tentativi ed errori è molto dispendioso ma aggiunge conoscenza sulla costruzione del processo a livello metodologico, tale conoscenza è importante tanto quanto la conoscenza che emerge dai risultati e l'azienda deve imparare a gestirla ed incamerarla perché non vada ad arricchire solo l'esperienza dell'analista che ha condotto il processo.

Alcuni di questi strumenti possono però offrire delle funzionalità che aiutano anche altri fasi del processo oltre alla fase di modellazione, ad esempio gli strumenti di visualizzazione dei dati sono molto utili nella fase di esplorazione e preparazione dei dati e nella fase di interpretazione dei risultati e trasferimento della conoscenza al

management. Gli strumenti di visualizzazione del processo sono altrettanto utili per tenere traccia dell'evolvere del processo.

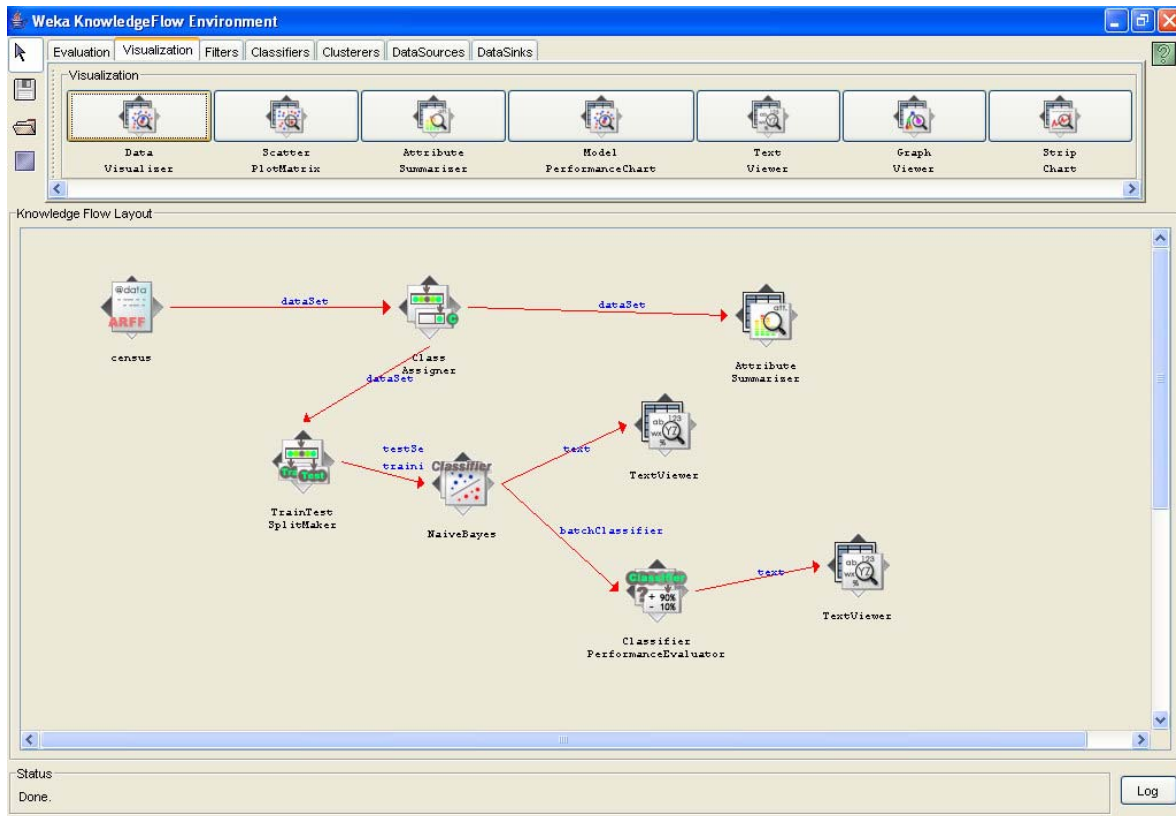


Figura 17: Interfaccia dell'ambiente di KnowledgeFlow fornito da Weka

La maggior parte degli strumenti commerciali per la costruzione di modelli sono praticamente equivalenti come funzionalità offerte, interfaccia grafica e costi.

Oltre alla scelta dello strumento di modellazione si è accennato alla decisione di come acquisire le conoscenze necessarie per sviluppare l'intero processo di data mining. Le possibili soluzioni spaziano dal completo sviluppo delle competenze all'interno dell'organizzazione al reperimento all'esterno di figure dedicate alla gestione del processo.

E' possibile rivolgersi a delle società di consulenza tecnologica per richiedere degli esperti di data mining per farli lavorare sui dati e sulla tecnologia presenti all'interno dell'organizzazione.

E' possibile anche rivolgersi alla società produttrice (o alla società distributrice) dello strumento di modellazione per beneficiare di un affiancamento di personale

specializzato per aiutare l'azienda a portare a regime il sistema e per insegnare all'analista come usare le varie funzionalità del nuovo strumento acquistato.

In generale la scelta tra acquisire del tutto le competenze all'esterno e tra tentare di sviluppare le competenze e la conoscenza all'interno è essenzialmente legato a quanto l'impresa ritiene centrale nel proprio business l'attività di business intelligence ed in particolare di data mining, se i processi di data mining vengono applicati una tantum e non riguardano aspetti del business ritenuti cruciali è giusto rivolgersi all'esterno se invece l'azienda vuole che la business intelligence e, in particolare, il data mining diventino una delle sue capacità e competenze distintive è indispensabile che gli esperti vengano creati all'interno. Quando un'impresa riesce a controllare e gestire il processo di data mining in tutti i suoi aspetti può beneficiare di tutti i vantaggi del processo essendo continuamente in fase di ascolto e di apprendimento nei confronti dei propri clienti.

Quale tra tutte queste alternative si adatta meglio all'Azienda del caso di studio?

Per quanto riguarda lo strumento di modellazione l'Azienda ha dimostrato nella gestione di altri aspetti tecnologici del progetto di CRM, come nel caso del datawarehouse, di preferire di mantenere internamente gli strumenti e la tecnologia necessaria. Per quanto riguarda lo sviluppo delle competenze informatiche, l'Azienda è spesso ricorsa all'acquisizione di persone qualificate da società di consulenza specializzate; nel caso della business intelligence però da una parte si sono acquisite competenze dall'esterno come gli esperti del datawarehouse, dall'altra si è anche voluta una figura di analista dei dati per il CRM che fosse interna. Il motivo di tale scelta è che, mentre gli aspetti tecnici possono essere gestiti anche da persone non interne all'azienda, gli aspetti legati alle informazioni e alla conoscenza devono essere gestiti all'interno.

Sulla base delle considerazioni su come è stato organizzato il lavoro fino a questo momento e sulla base di convincimenti personali dell'autore la soluzione più appropriata per l'Azienda è che venga acquistato uno strumento per la costruzione di modelli ad hoc rispetto ai problemi di business che di volta in volta si vuole analizzare e che si preveda una serie di affiancamenti necessari per la messa a regime dello strumento e per l'addestramento sull'uso dello strumento dell'analista o del team di

analisti che si occuperanno del processo di data mining. L'acquisto dello strumento però può essere anche secondario, si può infatti testare le capacità dell'ambiente attraverso un progetto pilota di data mining svolto con il supporto di esperti ed eventualmente in altre sedi dove ci sono più strumenti a disposizione. Ad esempio una realtà con la quale sarebbe molto interessante e stimolante collaborare è quella dei centri di data mining formati spesso da società produttori di software, società di consulenza e mondo universitario e della ricerca, una dei più importanti centri di questo genere in Italia è stato costituito da SAS con l'università SDA Bocconi di Milano, l'Università di Pavia, l'Università di Bologna ed altre.

Per quanto riguarda le competenze delle persone per la gestione del processo di data mining, le considerazioni fatte per l'acquisto del software rimangono valide.

Dal momento che l'Azienda in questione è alla sua prima esperienza con il data mining ma che ha la volontà di mantenere internamente le competenze e la conoscenza sulla gestione del processo di data mining, tutte le esperienze che dall'esterno potrebbero arricchire la conoscenza interna dovrebbero essere considerate molto seriamente in questo primo momento.

E' necessario che intorno ad una figura (o team) interna, l'analista, si crei una rete di persone, di società e di esperienze che supportano il processo e apportano la conoscenza e l'esperienza laddove ce ne sia bisogno.

L'ultima tipologia di risorsa che si prende qui in considerazione, senza entrare nel merito ma solo per motivi di completezza, è la risorsa finanziaria ed in particolare il fatto che debba esistere una proporzione definita a priori tra il budget che riguarda la costruzione dell'ambiente di *business intelligence* e il budget dell'intero progetto di CRM; un progetto di CRM dinamico infatti implica investimenti in tecnologia e competenze di business intelligence che crescono con il crescere del progetto stesso.

## **Lo stato dei dati**

La *business intelligence* si basa sull'estrazione di conoscenza dall'osservazione e dall'elaborazione automatica o semi-automatica dei dati, un ambiente di BI senza i dati, nel genere e nel numero necessari, non ha nessun senso; una volta stabiliti gli obiettivi

di business e di data mining bisogna indagare se l'azienda dispone dei dati giusti per raggiungere tali obiettivi e, nel caso non li disponga, quali sono gli interventi da compiere.

I dati sono caratterizzati dalla fonte di provenienze, dai possibili valori che contengono, dalla tipologia, dal livello di dettaglio rispetto al fatto che descrivono, dalla numerosità verticale o numero di record, dalla numerosità orizzontale o numero di colonne e dal livello di qualità.

Il sistema di *business intelligence* vecchio si basava essenzialmente sui dati provenienti dalle transazioni relativi agli acquisti nei negozi, dai dati che venivano memorizzati nella base di dati operativa ed organizzati secondo una logica multidimensionale.

Nell'analisi fatta nel capitolo 4 sullo stato dell'ambiente di business intelligence si è più volte fatto riferimento ai problemi che riguardano i dati a disposizione in Azienda: problemi di qualità dei dati dovuti alla fonte ed al meccanismo parzialmente automatizzato di monitoraggio, problemi di mancanza o scarsità di un certo tipo di dati come quelli anagrafici e socio-demografici, problema di mancanza di storia nei dati.

Alcuni di questi problemi sono in via di miglioramento e di soluzione, come nel caso della storicità non si può far altro che aspettare di accumularla nel tempo, mentre per altri è necessario un intervento piuttosto deciso.

La prima questione che si vuole affrontare è: *quanti dati servono per fare data mining?* In genere molti, milioni di record e varie centinaia di campi per ogni record ma anche questo aspetto può variare in base al tipo di business ed al livello di implementazione del processo in azienda. In generale è necessario che i dati siano numericamente e qualitativamente sufficienti a rappresentare la realtà che si vuole analizzare del business in tutti i suoi aspetti; avere milioni di dati scorretti o dove mancano le informazioni critiche che si vogliono indagare equivale a non avere dati. Il data mining però si basa sull'esplorazione di grandi quantità di dati per scoprire ed evidenziare relazioni e regole non note, se non c'è conoscenza implicita e nascosta nei dati perché i dati sono troppo pochi applicare il data mining non ha senso.

Infine, i dati devono contenere le informazioni necessarie che si vogliono indagare, rispetto agli obiettivi di data mining prefissati sarebbe utile avere:

- dati che riguardano la vita del cliente: di tipo anagrafico e demografico, variabili socio-economiche, comportamentali, sulle località di residenza.
- dati che riguardano il settore e il business (*prospect* anonimi): provenienti da indagini di mercato, ricerche, ecc.
- dati che riguardano l'azienda: budget, risultati economici, personale, organizzazione, ecc.
- dati che riguardano provenienti dai punti di contatto tra l'azienda e il cliente: call center, posta elettronica e tradizionale, sito web, negozio.
- dati sulle campagne marketing: descrizione delle campagne, risposte o non risposte dei contatti.
- dati di ritorno dalla *customer satisfaction audit*.

Per il momento l'unico canale di informazioni sul cliente è il negozio e i dati vengono memorizzati nella base di dati operativa, sarebbe interessante poter automatizzare altri punti di contatto già esistenti come il sito web per arricchire le informazioni sulla relazione con i propri clienti e non.

Le indagini dirette sui clienti, come quella condotta sulla soddisfazione del cliente, le informazioni sul settore, che riguardano il target di clienti potenziali o effettivi e la concorrenza, frutto di ricerche di mercato condotte dall'azienda o meno dovrebbero poter essere integrabili con i dati presenti nel database o nel datawarehouse.

Abbiamo già detto della difficoltà per l'Azienda di reperire dati sulla vita del cliente che non siano quelli basilari dell'anagrafica, una possibilità molto interessante è quella di comprare da altre banche dati informazioni sui propri clienti conosciuti, così come è possibile comprare nominativi di *prospect*, e cioè nominativi di persone che non risultano come clienti dell'Azienda ma che hanno le caratteristiche per diventarlo. Esistono anche banche dati pubbliche da cui reperire informazioni sulla geografia dei luoghi di residenza dei clienti o dei luoghi dove sono localizzati i negozi, informazioni di tipo socio-economico pubblicate dagli enti statali, ecc.

Infine l'Azienda si deve adoperare per la memorizzazione di tutte le informazioni che riguardano le campagne marketing dalla descrizione approfondita di tutte le sue caratteristiche alla descrizione dei comportamenti di risposta di ogni cliente alla campagna.

Dal momento che si parla di fonti di dati e canali molto diversi tra loro è lecito aspettarsi che sarà molto impegnativo rendere integrabili i dati che necessariamente saranno molto diversi tra di loro, i dati nel datawarehouse dovrebbero essere tutti riferibili al cliente ma laddove ci siano degli acquisti o delle interazioni con il cliente o *prospect* dovrebbe essere possibile scendere a livelli di dettaglio più alti come la transazione e il prodotto.

Come tradurre dati e informazioni sulla popolazione in generale e non direttamente riferibile ai clienti presenti nel database dell'Azienda è un'operazione piuttosto difficile ma indispensabile da compiere, un dato che non si riesce ad agganciare al cliente nel database è un dato che non può essere usato nel processo di data mining. Un problema correlato è quello della conoscenza implicita detenuta dalle persone esperte del business, tale conoscenza infatti deve essere opportunamente condivisa e interpretata dall'analista perché assuma un aspetto quantificabile e misurabile al fine di poter essere elaborata con il data mining. Nel dire questo si vuole affermare l'importanza che la conoscenza degli esperti di business riveste nella gestione del processo di data mining, deve essere innanzitutto comunicata e poi deve essere trasformata perché sia di aiuto nel processo di comprensione dei dati.

Le discussioni fatte fino a qui sui dati riguardano principalmente gli obiettivi di data mining di segmentazione della popolazione e hanno carattere generale, ci interessa ora prendere uno degli obiettivi di data mining, quello di previsione delle risposte alle campagne marketing, e scendere nel dettaglio dell'analisi di quali e quanti dati sono necessari.

Nel caso della costruzione di modelli predittivi delle risposte alle campagne è possibile determinare come devono essere fatte e quali siano le variabili necessarie per la costruzione del modello.

Un modello predittivo, indipendentemente dalla tecnica di data mining utilizzata per implementarlo, è basato su un modello creato sui dati che riguardano il passato dai quali emergono delle regole che possono diventare delle predizioni, chiamato *training model*, tale modello viene poi applicato su dati più recenti rispetto a quelli con cui si è costruito il modello per testare e valutare la vicinanza delle regole predette alla realtà.

I dati necessari per questo processo devono essere dati provenienti da campagne identiche o molto simili a quella attuale, tali dati devono contenere tutte le informazioni necessarie alla previsione e devono essere qualitativamente validi. Inoltre i dati usati per costruire il modello devono essere disponibili anche per la sua applicazione nella realtà, ad esempio, se per costruire il modello si è usata un'informazione che si conosce solo di alcuni clienti, il modello non potrà essere applicato su quei clienti di cui non si conosce l'informazione; questo è un problema che capita con i dati provenienti da interviste e indagini dirette fatte su campioni troppo piccoli o quando si sceglie di analizzare delle variabili di cui si conoscono i valori solo per pochi record.

Dal momento che il predittore del futuro è il passato è necessario valutare quanto la situazione presente si discosti da quella passata e quantificare l'impatto delle differenze tra il passato e il presente che possono verificarsi sul modello, tali differenze possono essere dovute a variabili esogene, cambiamenti sociali e politici intercorsi durante la vita dell'azienda ma anche cambiamenti avvenuti all'interno dell'azienda stessa. Ad esempio non è detto che esista in azienda una storia di campagne marketing identiche, possono essersi ripetute invece campagne marketing simili, bisogna saper identificare in che cosa consista la similitudine, isolare le caratteristiche che rendono l'attuale campagna diversa da quelle precedenti e infine quantificare l'impatto delle differenze sulle capacità del modello di produrre previsioni.

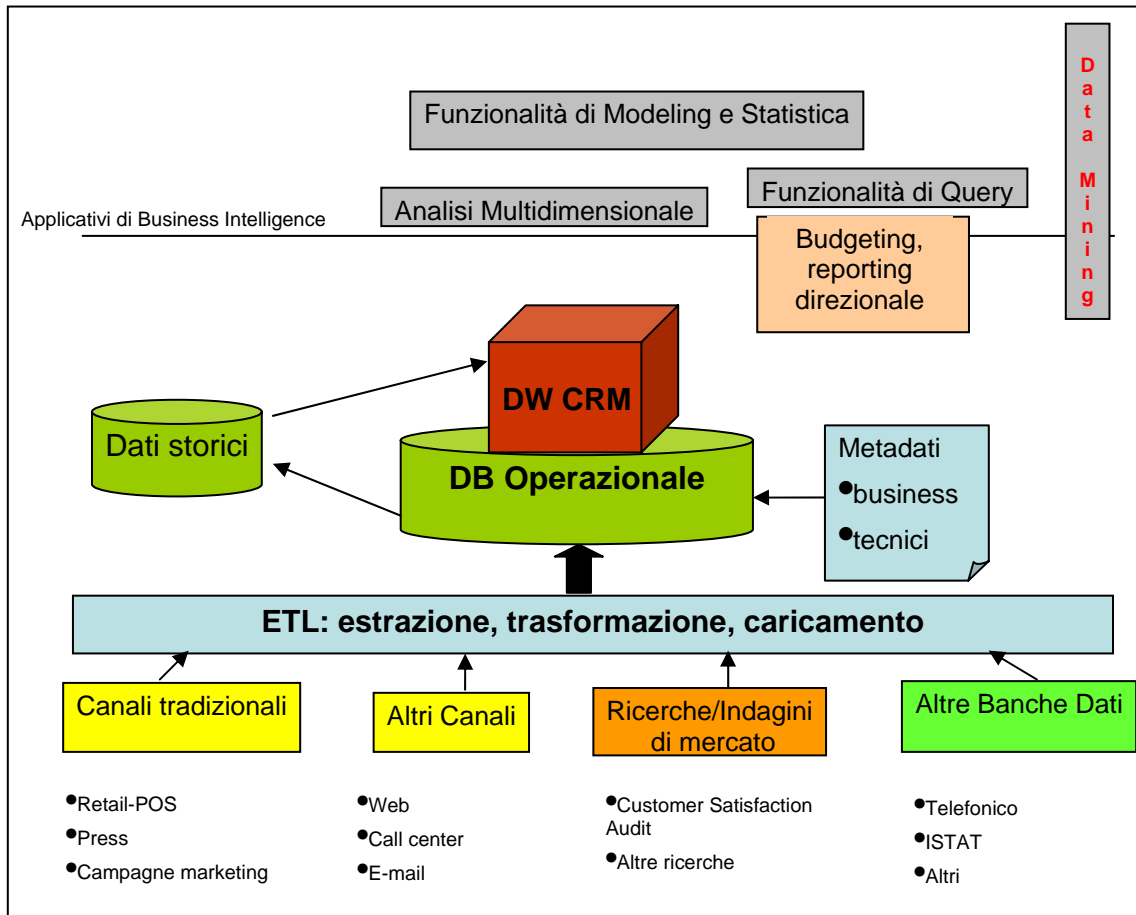
In questo contesto è importante coinvolgere gli esperti del business e dei dati che possono dare indicazioni sul passato, sui cambiamenti che si sono verificati, sulla presenza di una stagionalità del business che deve essere tenuta in considerazione nella scelta dell'intervallo di tempo oggetto di analisi, ecc.

Una campagna marketing è caratterizzata dall'aver un numero di contatti prestabilito solitamente su basi diverse da quelle del numero di record nel database ma deciso da considerazioni di budget, di marketing e di altro genere. Il numero dei clienti, o dei *prospect*, presenti nel database è a questo punto cruciale, se il numero non è sufficiente per costruire un predittore affidabile, costruire un predittore sarà inutile.

Inoltre una campagna marketing non sempre ha dei riscontri misurabili, che sicuramente non si troveranno laddove non è possibile sapere chi è stato contattato come nel caso di campagne marketing rivolte ad un pubblico indiscriminato attraverso i canali dei media.



L'analisi dei requisiti si conclude con uno schema di quella che potrebbe essere l'architettura di riferimento per l'ambiente di *business intelligence* che l'Azienda sta progettando di sviluppare.



**Figura 18: Architettura dell'ambiente di Business Intelligence**

## 5.4 Fasi dell'implementazione del processo di data mining

Nello studio di fattibilità per un ambiente di *business intelligence* fatto nel paragrafo precedente si è cercato di fare il punto della situazione rispetto a ciò che l'Azienda ha fatto e rispetto a ciò che l'Azienda deve fare per raggiungere i propri obiettivi; a completamento dello studio di fattibilità è interessante analizzare come dovrebbe essere condotto un processo di data mining contestualmente all'ambiente verso cui si vuole andare. Oltre all'analisi dei requisiti in base agli obiettivi è necessario tentare di immaginare come sarebbe un tipico processo di applicazione del data mining, da questa analisi emergerà con ancora più chiarezza se la direzione presa dall'Azienda è quella giusta e se le valutazioni sulle soluzioni sono valide o meno.

La simulazione di quello che potrebbe essere un processo di data mining da sviluppare nel nuovo ambiente di *business intelligence* verrà condotto seguendo il modello CRISP-DM (*Cross Industry Standard Process per il Data Mining*). Per l'approfondimento sulle caratteristiche del modello si rimanda all'appendice A.

Il modello CRISP-DM è un modello standard di processo di data mining, il modello individua sei fasi ordinate tra di loro, ciascuna di queste fasi è caratterizzata da uno o più obiettivi da raggiungere (*task*), da una serie di attività da svolgere per il raggiungimento di ogni obiettivo e una serie di documenti di output della fase che verranno presi in input nella fase successiva per il proseguimento del processo. Il processo ha un carattere ciclico, ogni fase è legata a quella successiva, con possibilità di loop, e le attività sono legate in un continuum interattivo e dinamico.

Le sei fasi consistono in:

1. Capire il business
2. Capire i dati
3. Preparare i dati
4. Modellazione
5. Valutazione
6. Trasferimento della conoscenza

**1. Capire il business:** la prima fase comprende tutte quelle attività necessarie all'analista che conduce il processo per capire gli obiettivi di business che si vogliono raggiungere e inquadrare questi obiettivi nel contesto aziendale ed organizzativo. Gli obiettivi di questa fase riguardano la comprensione del business, dell'azienda e molte delle attività da svolgere in questa fase sono già state descritte ed affrontate nei paragrafi precedenti di questo capitolo e nel capitolo 4.

#### *Determinare gli obiettivi di business*

L'analista deve essere in grado di capire e recepire quelli che sono gli obiettivi di business espressi dall'Azienda, spesso infatti in azienda non è veramente del tutto chiaro quello che si vuole fare. L'Analista e il Manager di progetto interpretano le aspettative dell'Azienda nei confronti del data mining e analizzano la fattibilità e sensatezza delle richieste pervenute.

Ipotizziamo che da tale riflessione tra l'analista e il manager di progetto emerga la definizione che l'obiettivo di business per l'Azienda sia di *aumentare l'impatto positivo sulle vendite dell'invio dei cataloghi pubblicitari della nuova collezione*.

Definito l'obiettivo l'analista deve indagare tutti gli aspetti che riguardano il problema di business, il risultato dell'indagine è l'identificazione di tutte le risorse a disposizione del processo attuale e la comprensione di come il problema di business veniva gestito precedentemente.

La soluzione adottata attualmente in Azienda è che la spedizione dei cataloghi è diventato una campagna marketing gestita dall'ufficio CRM dal momento della costituzione del progetto. Per ogni collezione (autunno/inverno, primavera/estate) vengono spediti 1-2 cataloghi per tipologia di negozio/prodotto. Per il momento il numero di cataloghi disponibile per la spedizione è determinato da una decisione che coinvolge solo in parte l'ufficio CRM. Il numero di cataloghi è un input del processo, così come lo è il numero di clienti nel database; a partire da questi due input viene elaborata una lista di clienti ordinati per ammontare speso in ogni tipologia di negozio. La lista viene estrapolata in base ai requisiti di consenso esplicito da parte del cliente alla ricezione di materiale pubblicitario per posta, al fatto che la persona sia effettivamente un cliente cioè che abbia effettuato un acquisto recentemente in uno dei negozi che fa parte della tipologia di negozio

pubblicizzata dal catalogo. Infine di ogni cliente selezionato devono essere conosciuti con esattezza: *titolo, nome, cognome e indirizzo completo e postalizzabile*. La lista di clienti a cui è stato spedito il catalogo viene memorizzata a sistema insieme a tutte le informazioni che riguardano la campagna (tipo di catalogo, data di invio, data di ricezione presunta, numero di cataloghi inviati, ecc.). Nelle settimane successive all'invio viene fatta una raccolta dei ritorni postali causati da problemi di vario genere di correttezza nell'indirizzo, per ogni catalogo non recapitato viene modificato il profilo del cliente corrispondente e viene tenuta traccia del numero di cataloghi non recapitati, viene aggiunta l'informazione alla campagna del numero preciso di cataloghi arrivati a destinazione.

Inoltre viene applicato anche un meccanismo di "destinatari civetta", viene mandato il catalogo anche ad un certo numero di persone legate all'Azienda, solitamente dipendenti, che hanno il compito di comunicare all'ufficio se e quando hanno ricevuto il catalogo, sulla base di queste segnalazioni si esegue una stima dell'effettivo numero di cataloghi ricevuti. Per quanto riguarda come l'Azienda fino a questo momento abbia gestito la misurazione dell'impatto della spedizione dei cataloghi sul comportamento di acquisto si è proceduto in maniera molto simile a come descritto nel paragrafo 4.5 per la misurazione dei risultati dei *trunk show*: l'analista con lo strumento di analisi *BusinessObjects* recupera la lista di clienti inseriti nella campagna e osserva il loro comportamento di acquisto nelle cinque settimane successive alla ricezione presunta del catalogo. Il documento di reporting indica i risultati di sintesi dell'analisi descrittiva del comportamento dei clienti che hanno ricevuto il catalogo.

#### *Conoscere la situazione aziendale in cui si stabilisce il processo*

E' utile condurre un'analisi approfondita di quali sono le risorse effettivamente dedicate al processo in termini di strumenti software e hardware e fonti dei dati ma anche in termini di persone. Le persone coinvolte nel processo sono quelle rappresentate in figura 16 dell'organigramma delle persone e dei ruoli coinvolti nel progetto, gli esperti del business sono il manager di progetto CRM, i manager *Retail* e i manager *Press*, gli esperti delle fonti dei dati sono i responsabili del software di

gestione del database clienti e del POS, i responsabili del *datawarehouse*, l'analista dei dati.

Gli strumenti software a disposizione sono *BusinessObjects*, funzionalità di *query* per accedere direttamente al database e uno strumento di modellazione.

Per quanto riguarda i vincoli tecnologici: i dati provengono principalmente dal *datawarehouse* ma potrebbero anche provenire da altre fonti non ancora integrate con il *datawarehouse*. Se il *datawarehouse* è già stato implementato su un ambiente dedicato non ci sono vincoli all'accesso ai dati, altrimenti si dovrà considerare l'allungamento dei tempi dovuto all'accesso alla base di dati operativa. I dati relativi alle anagrafiche clienti vengono periodicamente normalizzati da una società esterna, l'intervallo di tempo tra la consegna dei dati alla società e la ricezione dei dati è un momento in cui sarebbe meglio non compiere azioni sui clienti.

#### *Determinare gli obiettivi di data mining*

L'obiettivo di business deve essere tradotto in termini tecnici e potrebbe essere definito come: *prevedere la risposta dei clienti o prospect alla ricezione di un catalogo pubblicitario di una collezione*. Questo implica, in termini di data mining: *classificare i clienti in base alla loro propensione di acquisto successivamente alla ricezione del catalogo*.

Preliminarmente a questo obiettivo sarebbe opportuno segmentare attraverso tecniche di *clustering* la popolazione dei clienti in base alle caratteristiche anagrafiche e di comportamenti di acquisto per individuare dei gruppi di clienti simili, ma questo è un altro problema di data mining che dovrebbe essere gestito con un altro processo.

Contestualmente alla determinazione degli obiettivi di data mining deve essere anche definito con quali metodi si valuterà il successo del modello, dal momento che parliamo di previsione uno dei criteri di successo è l'accuratezza della previsione, un altro aspetto che può essere considerato di successo è la facilità di lettura dei risultati del modello.

### *Produrre un piano ben definito del progetto*

Infine è necessario pianificare tutte le fasi del processo e discuterne la tempistica con le persone dell'Azienda coinvolte. Il piano oltre a specificare le attività e le operazioni in ordine cronologico deve anche definire le tecnologie e gli strumenti a disposizione al momento dell'avvio del processo sulla base delle considerazioni fatte nell'analisi dell'organizzazione e nell'individuazione degli obiettivi di business e di data mining.

**2. Capire i dati:** questa fase consiste nell'esplorazione dei dati a disposizione nell'organizzazione ed eventualmente recuperabili da fonti esterne. La comprensione dei dati è condizione necessaria per la costituzione di un processo di estrazione di conoscenza dai dati. In questa fase è molto importante, come in quella successiva, la collaborazione e l'interattività tra l'analista e gli esperti dei dati e del business dell'organizzazione.

### *Collezionare i dati iniziali*

L'attività principale è quella di pianificare quali sono le variabili che servono per l'analisi e individuare quali sono le fonti da cui estrarre tali variabili.

Per l'acquisizione dei dati si devono mettere a punto le procedure di accesso alle diverse fonti che verranno eseguite con gli strumenti di accesso e caricamento dei dati che sono stati identificati nella fase precedente. Ad esempio i dati che riguardano il database clienti e le vendite associate ai clienti sono reperibili o direttamente dalla base di dati operativa sotto forma di tabelle da collegare successivamente oppure possono essere acceduti attraverso l'interfaccia di interrogazione del *datawarehouse* fornita da *BusinessObjects*, accedendo i dati in questa maniera si mantengono le logiche cliente-centriche già implementate nel *datawarehouse*.

Se l'Azienda ha acquisito nuovi dati sui clienti o sui negozi da banche dati esterne o da fonti interne mai utilizzate prima e tali nuove fonti non sono ancora state integrate nel *datawarehouse* sarà necessario introdurre anche queste nuove fonti e predisporre una procedura di accesso, trasformazione e caricamento dei dati.

Le variabili che si vogliono comprendere nell'analisi sono tutte quelle disponibili sul cliente, dalle variabili che riguardano il cliente come individuo, alle variabili calcolate o derivate che descrivano i comportamenti di acquisto del cliente, variabili sulle risposte precedenti alle passate campagne marketing, con particolare riferimento all'invio dei cataloghi e tutte le informazioni che riguardano le campagne marketing.

La fonte principale è il *datawarehouse* sia per quanto riguarda il cliente sia riguardo al suo comportamento di acquisto e infine le campagne e il negozio. Sarebbe interessante avere riferimenti geo-politici sulla località dove è situato ogni negozio, riferimenti socio-demografici sulle località di residenza dei clienti, informazioni sulla vita del cliente, reddito, professione, famiglia, ecc.

#### *Descrivere i dati*

Una volta acquisiti i dati si procede alla descrizione di questi in termini di quantità e tipo di dato.

#### *Esplorare i dati*

L'esplorazione dei dati serve per approfondire la conoscenza dei dati acquisita fino a questo momento; l'esplorazione viene condotta con l'aiuto di strumenti di *query*, visualizzazione e *reporting*. Se lo strumento di data mining scelto dall'Azienda offre funzionalità di visualizzazione dei dati di input sarà molto utile avvalersi di tali strumenti in questa fase e nel caso non si possedessero si può decidere di comprare uno strumento dedicato di visualizzazione. L'Azienda già dispone di strumenti di *query* e *reporting*, *BusinessObjects*, con tale strumento si può condurre abbastanza agevolmente l'analisi delle proprietà degli attributi più significativi, ad esempio calcolando le misure statistiche di base.

L'esplorazione può essere condotta anche al fine di verificare alcune ipotesi che sono state fatte sui dati o sul business, ad esempio il manager *Retail* può aver comunicato all'analista che la maggior parte dei clienti che fanno un acquisto nei negozi di Milano il sabato pomeriggio è in un'età inferiore ai 25 anni, se il target giovani sotto 25 anni è di particolare interesse nelle analisi sarà utile approfondire fin da questo momento del processo se tale ipotesi ha un qualche fondamento.

Alcune considerazioni che vengono dagli esperti di business in realtà sono ipotesi che si basano sull'intuito e l'esperienza sul campo, può essere importante verificare se i dati confermano tali intuizioni, specialmente quando su quelle assunzioni si basa buona parte del processo di analisi.

#### *Verificare la qualità dei dati*

Come già ripetuto varie volte la problematica relativa alla qualità effettiva dei dati a disposizione dell'organizzazione è fondamentale e vitale per il processo. Nel caso della nostra Azienda al momento attuale ci sono ancora dei problemi e fino a quando ci sarà un problema di qualità dei dati sarà necessario corredare qualsiasi risultato emerso dall'analisi dei dati con il relativo margine di errore dovuto allo stato dei dati.

Per far questo devono essere controllati lo stato di tutti i campi e i valori che ogni campo può assumere, per ogni campo deve essere calcolata la percentuale del numero di *missing values*, gli eventuali errori di battitura dei campi stringa, gli eventuali valori fuori intervallo o anomali, ecc. Gli attributi che hanno percentuali di errore troppo alti devono essere scartati dall'analisi.

Un discorso particolare deve essere fatto per quanto riguarda la correttezza dei campi dell'indirizzo postale, la correttezza in questo caso è necessaria anche per le azioni di postalizzazione successive all'analisi. Acquisire e mantenere gli indirizzi postali è un problema molto importante e critico tanto che l'Azienda ha deciso di esternalizzare l'attività di pulizia e manutenzione dei dati necessari alla postalizzazione, *Nome, Cognome e indirizzo completo*. Questa peculiarità è stata evidenziata nella fase di conoscenza dell'organizzazione ed è stata considerata nella pianificazione delle attività del processo perché pone dei vincoli di tempo.

**3. Preparare i dati:** con questa fase si entra sempre più nel merito della parte tecnica del processo, l'obiettivo finale di questa fase è di produrre l'insieme dei dati nella forma che lo strumento di modellazione richiede in input. La fase di preparazione dei dati dovrebbe essere una fase tecnica e alcune delle sue attività possono essere automatizzate; il suo successo dipende fortemente dalle due fasi precedenti dove sono state affrontate tutte le problematiche concettuali e pratiche del



business e dei dati. In realtà in questa fase viene effettuato una ulteriore verifica dello stato dei dati, durante il lavoro di pulizia e costruzione dei dati può emergere che le variabili scelte non sono quelle adatte e appropriate per l'analisi, in tal caso sarà necessario tornare sui propri passi e ripercorrere le due fasi del processo precedenti.

### *Selezionare i dati*

L'obiettivo è quello di scegliere quali dati inserire nell'analisi perché utili concettualmente, validi qualitativamente e disponibili coerentemente con i vincoli di cui si è già detto.

I dati selezionati provenienti dal *datawarehouse* riguardano i clienti che hanno un nome, cognome e indirizzo validi come dovrebbe emergere dall'attività di normalizzazione, di questi clienti si selezionano i campi contenenti i dati anagrafici a disposizione (città di residenza, età, sesso), i dati di vendita negli ultimi due anni (data primo acquisto, data ultimo acquisto, numero di scontrini acquistati, ammontare speso medio, ammontare speso complessivo, distanza media tra un acquisto ed un altro). Infine i dati che riguardano i contatti tra l'Azienda e il cliente avvenuti tramite le campagne marketing (numero di campagne in cui il cliente è stato coinvolto, indicatore sintetico di risposta per ogni campagna), e in particolare i dati sui cataloghi ricevuti (numero di cataloghi ricevuti, tipologia di catalogo, stagione del catalogo e i dati sul comportamento di acquisto successivo alla ricezione dei cataloghi, ad esempio effettuato l'acquisto, tipologia negozio, tipologia prodotto, indicatori comportamenti di acquisto riferiti al periodo successivo alla ricezione del catalogo).

Per quanto riguarda altre fonti di dati, se non sono già state integrate con il *datawarehouse*, si vogliono aggiungere gli attributi sulla località dove si trova il negozio, come ampiezza del comune, macro area geografica, caratteristiche dell'area e della località (zona industriale, commerciale, turistica, ecc.). Ed infine sarebbe utile arricchire i dati con attributi sulle località di residenza del cliente come per il negozio e scendere più nel particolare, ad esempio sfruttando le informazioni contenute nel codice postale, aggiungere informazioni come reddito medio pro-capite della zona, grado di istruzione, ecc.

### *Pulizia dei dati*

I dati devono essere adattati al livello richiesto dalla tecniche di analisi scelte. Durante le attività svolte per la pulizia dei dati è possibile che emergano dei vincoli o dei problemi con i dati scelti e sarà quindi necessario dover tornare alla fase precedenti per riconsiderare i criteri di scelta delle variabili.

### *Costruire i dati*

Molto spesso i dati che vengono scelti per l'analisi sono dati costruiti a partire da uno o più dati appartenenti allo stesso record. Gli *attributi costruiti* servono per esprimere dei fatti che non emergerebbero con abbastanza enfasi dai dati estratti direttamente dal database o dalle altre fonti dei dati.

### *Integrare i dati*

Come già detto più volte i dati collezionati fin qui devono essere infine integrati in un'unica tabella, questa operazione può rivelarsi onerosa e difficile e può incontrare degli ostacoli che obbligano a tornare alla fase precedente e riconsiderare la scelta degli attributi.

Nel caso della nostra Azienda devono essere collegate la tabella proveniente dal datawarehouse dove ogni record corrisponde ad un cliente e la chiave del record è il codice cliente con le tabelle contenenti le informazioni geografiche dove la chiave è l'identificativo della località.

### *Formattare i dati*

Infine la tabella costruita con gli attributi scelti deve essere convertita nel formato che richiede lo strumento di modellazione. Il tipo di formato richiesto cambia da strumento a strumento ma in generale sarà un formato che comprende gli attributi divisi in colonne da un separatore e una parte del file deve essere dedicata ai metadati, cioè a quelle informazioni che descrivono i dati del file, di solito riportano informazioni come il nome dell'attributo, il tipo dell'attributo, il numero di caratteri se è una stringa, il limite superiore ed inferiore se è un numerico, gli estremi degli intervalli se è categorico, ecc.

**4. Modellazione:** questa fase rappresenta il nocciolo operativo del processo e per questo motivo per data mining spesso si intende solo il contenuto e le operazioni che si svolgono nella fase di modellazione. La modellazione ha un contenuto tecnico e dovrebbe essere condotta da specialisti, la modellazione può essere però in parte automatizzata a seconda del tipo di strumento che si usa. Come detto nel paragrafo 5.3 nell'analisi degli strumenti di data mining in commercio, esistono strumenti che forniscono dei modelli precostituiti che si adattano ad una serie di problemi di data mining e di contesti di business ma esistono anche strumenti che permettono di costruire il modello ad hoc sulla base delle caratteristiche dell'azienda. Assumiamo invece che l'Azienda si doti di uno strumento per costruire modelli e vediamo quali sono gli obiettivi di questa fase.

#### *Selezionare la tecnica di modellazione*

E' necessario soffermarsi sul fatto che lo strumento di data mining scelto offrirà solo un sottoinsieme di tecniche di modellazione, implementate a loro volta da un sottoinsieme di algoritmi tra tutti gli algoritmi ideati per ogni tecnica.

Tra le tecniche offerte dallo strumento si dovrà scegliere quella più adatta al problema. In realtà il processo dovrebbe essere inverso, cioè lo strumento dovrebbe essere scelto in base ai vincoli tecnologici ed organizzativi presenti in azienda, ed in base alla gamma di tecniche offerte dai vari strumenti.

In generale, la tecnica degli alberi decisionali è supportata da tutti gli strumenti commerciali più diffusi ed in questo caso sceglieremo proprio gli alberi decisionali per il nostro problema di classificazione. Dal momento che ogni tecnica può essere implementata a sua volta da più di un algoritmo diverso anche la stessa tecnica può presentare requisiti diversi a seconda dello strumento scelto, ad esempio la maggior parte degli strumenti propongono una versione di albero decisionale che richiede che gli attributi di input siano tutti categorici, questo rappresenta un vincolo piuttosto restrittivo! Se solo a questo punto del processo si viene a conoscenza di vincoli di questo genere sarà necessario tornare indietro alla fase di preparazione dei dati.

### *Generare una procedura o un meccanismo con cui testare il modello*

Prima ancora di costruire il modello sarà necessario definire la procedura o il meccanismo con cui si testerà la qualità e la validità del modello. Nel caso di problemi di data mining supervisionato come il nostro si usa il tasso di errore come misura della qualità del modello.

### *Costruire in modello*

Si sottopone l'insieme dei dati generato allo strumento di modellazione. Lo strumento richiede di assegnare dei valori ad una serie di parametri, la scelta di tali valori è critica e deve essere documentato il motivo della scelta.

Una volta sottoposti i dati e settati i parametri lo strumento produce dei risultati, nel caso degli alberi i risultati possono essere visualizzati o come delle regole tipo  $A \wedge B \rightarrow C$  oppure possono essere visualizzati a forma di albero.

L'output di questo *task* è costituito dai risultati presentati in una forma appropriata e dalla descrizione del modello, della sua accuratezza e robustezza e del processo di costruzione.

Solitamente non viene costruito un solo modello, ne può essere costruito più di uno cambiando tecnica e algoritmo e questo implica che tutte le fasi vengano ripercorse, mantenendo invece i risultati conseguiti con il processo fino a questo punto si può produrre più di un modello semplicemente cambiando i valori dei parametri.

### *Migliorare il modello (valutazioni tecniche)*

I risultati del modello devono essere analizzati per verificare che siano coerenti con gli obiettivi di data mining prefissati.

Esistono vari metodi per valutare il livello di predizione del modello, esistono anche diverse strategie per valutare i risultati ottenuti, si può scegliere tra la metodologia che consiste nel dividere l'insieme dei dati in un insieme che serve per allenare il modello e un altro insieme per testare i suoi risultati o si possono scegliere metodi diversi come quello di *cross-validation*.

Nel caso si fosse costruito più di un modello per scegliere qual è il migliore si può usare come misura il *lift*.

Per migliorare il modello si deve anche analizzare i risultati nei termini di business, per controllare che i risultati siano in accordo o meno con le assunzioni e la conoscenza esistente in azienda e che i risultati siano comprensibili e abbiano un senso per l'organizzazione.

**5. Valutazione:** in questa fase per valutazione si intende quella che riguarda il livello di raggiungimento degli obiettivi di business mentre tutte le considerazioni sulla qualità e correttezza tecnica del modello sono già state fatte nella fase precedente. Il modello potrebbe avere dato dei risultati molto accurati e che ben predicono i fatti del business ma potrebbe non aver messo in luce niente di nuovo rispetto al fatto che si voleva indagare.

L'output totale del processo infatti non è solo il risultato del modello ma è la combinazione di questo con tutto ciò che è emerso di nuovo durante la vita del processo. L'output di un processo di data mining può essere ad esempio una nuova domanda, un nuovo dubbio o può essere "semplicemente" la constatazione che i dati non sono adeguati all'analisi e servono nuovi interventi per raccoglierne di nuovi e con una qualità più alta.

Nel nostro caso i criteri di successo del modello sono la capacità di predire quali clienti rispondono positivamente alla ricezione del catalogo pubblicitario individuando le variabili chiave che caratterizzano queste persone. Dal processo di analisi l'Azienda si aspetta un modello predittivo efficace ma non solo; per l'Azienda il processo che stiamo descrivendo rappresenta uno dei primi approcci al data mining e quindi la valutazione del modello e del processo assumono contenuti più ampi, l'Azienda sta valutando se il data mining può effettivamente aiutare il business.

La valutazione del processo dipende fortemente da quanto utili sono le informazioni estratte per il business in proporzione a quanto lavoro è stato fatto per ottenere tali informazioni.

L'utilità delle informazioni non implica che la conoscenza debba riguardare esclusivamente l'obiettivo iniziale di data mining, l'Azienda si aspetta che verrà aggiunta della conoscenza e in questa fase di test delle potenzialità del sistema e

dell'ambiente è ben disposta ad accogliere tutti i tipi di risultati che possono arrivare a patto che siano comprensibili e di reale supporto alle attività dei manager.

Ad esempio dal modello di previsione delle risposte alla ricezione del catalogo può emergere che i cataloghi non vengono spediti nel momento giusto della stagione, questa informazione esula dall'obiettivo di data mining prefissato ma è senza dubbio molto utile.

#### *Valutazione dei risultati*

In questa fase si vuole valutare l'impatto dei risultati di data mining rispetto agli obiettivi di business, la valutazione passa attraverso le attività di comprensione dei risultati di data mining, interpretazione dei risultati in termini di applicabilità, effettivo arricchimento della conoscenza grazie ai risultati e all'intero processo.

E' stato più volte ripetuto come il processo di data mining apporti conoscenza, in ognuna delle sue fasi, scoprendo e mettendo in luce aspetti del business, dei dati e dell'organizzazione critici, in questa fase di valutazione devono essere considerati anche tutti i risultati del processo che non sono direttamente correlabili all'output della modellazione.

I risultati devono essere valutati anche rispetto ai risultati che si ottenevano con il precedente metodo di scelta per l'invio dei cataloghi; che consisteva nell'invio del catalogo ai clienti con la spesa più alta relativa ad un certo periodo di acquisto. Se i risultati ottenuti perfezionano solo di poco il metodo precedente la valutazione del processo sarà ovviamente non positiva.

#### *Rivisitazione del processo*

A questo punto la riflessione sulle fasi del processo e la rivisitazione delle attività e delle motivazioni alla base di ogni scelta costituiscono un'analisi della qualità del processo nel suo insieme.

L'output di questo *task* è un documento che sintetizza il processo svolto e mette in risalto le eventuali disattenzioni o errori commessi e che non devono essere ripetuti.

### *Definizione dei prossimi passi*

Sulla base delle valutazioni tecniche e di business si deve decidere come procedere d'ora in poi, se procedere alla fase successiva o se ritornare su alcune delle fasi precedenti e farle diversamente.

E' fondamentale per questa decisione quanto si ritiene che i risultati possano essere usati per supportare e guidare le attività di business o se ci sono ancora delle potenzialità inesprese nei dati e nel modello che è bene far emergere.

**7. Deployment o trasferimento della conoscenza:** la conoscenza emersa dal processo deve essere adeguatamente comunicata a chi ricopre il ruolo giusto per tradurre tale conoscenza in azioni. Il trasferimento della conoscenza può essere agevole e lineare ma può anche essere molto complicato. Il processo di data mining affronta problemi sia di tipo aziendale che di tipo tecnologico ed informatico, per gestire e sviluppare un processo di data mining devono essere coinvolte persone con competenze molto diverse tra di loro, il processo è gestito e sviluppato da esperti con conoscenze e competenze specialistiche di data mining, database ed organizzazione. La conoscenza che emerge da un processo così complesso è necessariamente non sempre accessibile e comprensibile da tutti, per leggere i risultati del modello servono conoscenze di statistica e di informatica che solitamente gli esperti di business non hanno, per individuare i fattori critici di successo serve una conoscenza del business e dell'organizzazione che solitamente gli esperti di data mining non hanno. Lo staff dedicato al progetto dovrebbe includere una figura che funga da interfaccia tra i due mondi e che abbia le competenze distintive di entrambe le categorie, tale figura di solito è interna all'organizzazione perché deve conoscere le peculiarità dell'azienda, del settore in cui opera e dei dati e delle tecnologie in uso, deve però anche avere una formazione scientifica di tipo informatico o statistico che le permette di seguire e comprendere tutte le fasi tecniche del processo.

### *Pianificare il trasferimento della conoscenza*

L'Azienda del nostro caso di studio ha sviluppato e gestito il progetto con la presenza della figura dell'analista, l'analista è una figura interna con competenze

informatiche di database, datawarehousing e analisi dei dati e con conoscenze del contesto di business in cui l'Azienda opera e soprattutto conoscenze dell'organizzazione acquisite lavorando in azienda nell'area del marketing. Una figura come l'analista è la figura ideale per gestire al meglio tutto il processo e soprattutto la fase di *deployment*, quando manca questa figura manca l'anello di congiunzione tra gli attori con conoscenze manageriali e gli informatici o gli statistici che seguono la parte tecnica del processo di data mining e ciò significa che non c'è nessuno capace di tradurre i risultati tecnici in informazioni comprensibili per l'organizzazione.

Dal momento che i risultati del processo di data mining giocano un ruolo strategico nel supportare le decisioni dei manager, l'Azienda ha bisogno di assicurarsi sul modo in cui i risultati sono stati prodotti, è necessario cioè che i risultati possano essere adeguatamente giustificati e motivati. Il ruolo dell'analista, se identificato in una persona interna e di fiducia dell'azienda funge anche da garante sulla qualità e validità dei risultati portati dall'intero processo.

#### *Pianificare il monitoraggio e la manutenzione del processo*

Il monitoraggio e la manutenzione del processo sono attività importantissime se si ha intenzione di ripetere il processo più volte e di inserirlo tra le attività caratteristiche del business.

La pianificazione deve prendere in considerazione gli eventuali cambiamenti che possono avvenire nel tempo a cui il processo si dovrà adattare, gli intervalli di tempo con cui si dovrà ripetere il processo, ecc.

#### *Produrre il documento finale*

Alla fine del processo il manager di progetto e il suo team prepara un *report* finale in cui descrive i punti più interessanti del processo, i risultati rispetto agli obiettivi e i risultati emersi inaspettatamente. Il tipo di *report* dipende da come si è gestito il trasferimento della conoscenza.



*Review del processo*

A questo punto si hanno tutti gli elementi per fare un bilancio di ciò che è andato bene e ciò che è andato male e per individuare ciò che deve e può essere migliorato.

## 6. Conclusioni

Con questo lavoro si è cercato di raccontare in maniera critica e analitica il percorso intrapreso da un'azienda nella costruzione di un ambiente di *business intelligence*; nel corso della trattazione l'interesse principale è stato rivolto al rapporto che intercorre tra l'organizzazione e le tecnologie informatiche di *business intelligence*.

Il caso di studio scelto è particolarmente interessante perché riguarda un'azienda il cui business e le cui competenze esulano completamente dall'*information technology* ma che ha capito di potersi e doversi servire delle nuove tecnologie informatiche per acquisire un vantaggio competitivo rispetto ai concorrenti e richiesto dal momento storico ed economico.

Disporre di un ambiente di *business intelligence* come quello descritto fin qui, saperlo usare e saperne sfruttare le grandi potenzialità, fa la differenza tra un'azienda e le altre che operano nello stesso settore e questa affermazione vale indipendentemente dal settore.

Nella tesi si è analizzato uno dei campi di applicazione dei sistemi della business intelligence, il CRM, che coinvolge principalmente il marketing e le vendite; è importante puntualizzare però che i sistemi di *business intelligence* possono e dovrebbero essere diffusi in tutti le aree dell'organizzazione.

Tale precisazione è essenziale perché l'applicazione di questi sistemi avviene solitamente solo nelle aree funzionali che ne hanno sponsorizzato l'introduzione iniziale, come succede anche nel nostro caso di studio, e perché spesso essi vengono applicati solo su un sottoinsieme dei problemi decisionali di competenza della stessa area; in generale però non vi è a priori nessuna ragione o vincolo perché le tecnologie di business intelligence non possano essere applicate all'azienda nella sua totalità, lungo tutta la catena del valore e persino nella rete del valore, estendendo l'accesso alla filiera dei clienti o ai fornitori e partner.

La piattaforma informatica di supporto a questi sistemi di management dovrebbe e potrebbe diventare una piattaforma standard utilizzata a livello aziendale, seguendo un naturale percorso di diffusione ed estensione interfunzionale o interdivisionale. Al riguardo giocano sicuramente un ruolo decisivo la definizione di una strategia di un

ambiente di business intelligence a livello aziendale, la diffusione dell'esperienza e delle competenze costruite nei primi progetti, la sofisticazione della piattaforma tecnologica, i risultati e la soddisfazione generati, tutti fattori che possono ampliare la portata dell'ambiente sia in termini di numero di fatti e di fenomeni aziendali di competenza di molteplici unità organizzative, sia in termini di maggior livello di dettaglio raggiungibile nei dati disponibili.

Nel caso dell'esperienza dell'Azienda, il sistema di *business intelligence* è stata introdotto a sostegno e completamento del progetto pilota di CRM; il progetto è passato attraverso la fase iniziale pilota per poi diventare un progetto consolidato, l'ambiente di *business intelligence* ha seguito dinamicamente le fasi dello sviluppo del progetto.

Come già detto la strategia di CRM, per sua natura, per essere implementata ha bisogno del supporto di un ambiente di *business intelligence*; un sistema di business intelligence supporta la strategia e crea nuove opportunità di sviluppo per il progetto stesso.

*Lo sviluppo del progetto impone uno sviluppo conseguente della tecnologia:* quando il progetto di CRM è giunto alla fase di consolidamento, si è iniziato a valutare l'opportunità dell'estensione dell'ambiente con funzionalità di data mining perché dopo aver sperimentato e beneficiato della potenza espressiva e descrittiva delle analisi OLAP sono nate nuove domande e nuove esigenze informative a cui però l'analisi OLAP non è in grado di rispondere soddisfacentemente.

*La tecnologia propone nuove opportunità di sviluppo:* la tecnologia del data mining implica, ancora prima di essere implementata, una serie di riflessioni sul business e sull'organizzazione e comporta spesso interventi adattivi e cambiamenti delle tecnologie, dei ruoli e delle competenze presenti in azienda.

Un'impresa proattiva deve investire pesantemente nella costruzione di un ambiente di *business intelligence* che la supporti nell'applicazione della strategia di apprendimento e cambiamento continuo e che la prepari ad affrontare le nuove opportunità.

Lungo tutto il corso della trattazione si è voluto mettere in evidenza, con esempi e riflessioni, il fatto che per affrontare cambiamenti così importanti e pervasivi come l'introduzione di un ambiente di business intelligence o l'estensione dell'ambiente con nuove e più sofisticate funzionalità, è necessario svolgere un'attività strutturata e

razionale di riflessione sulle strategie e sugli obiettivi da parte del management guidata e supportata da esperti di *business intelligence* e in base ai risultati di questa fase procedere ad un'analisi dei requisiti organizzativi e tecnologici necessari per tradurre in pratica gli obiettivi del management.

I sistemi di *business intelligence* incorporano tecnologie informatiche molto variegate, che comprendono tutti gli applicativi utilizzati all'interno dell'organizzazione (DBMS, applicativi per la produzione, comunicazione tra le aree, ecc.) e all'esterno (canali di contatto, negozio, forza vendita, ecc.); l'esperto di business intelligence deve avere una formazione informatica avanzata che gli permette di capire e seguire gli aspetti essenziali di queste tecnologie ed in particolare deve avere competenze specifiche di basi di dati, datawarehousing e di data mining, competenze che a loro volta implicano una conoscenza approfondita dei dati presenti in azienda e conoscenze di statistica.

Più volte si è fatto riferimento al fatto che un sistema di *business intelligence* produce risultati utili al management e che per tradurre l'output del sistema in informazioni fruibili a livello direzionale è necessario che esista una figura, l'esperto di *business intelligence* o analista, che abbia conoscenze e competenze informatiche molto approfondite ma anche conoscenze del business, di economia aziendale e capacità di comunicare e relazionarsi con il management spesso non dotato di una formazione tecnico-scientifica.

La figura dell'esperto di *business intelligence* ha un ruolo operativo nell'organizzazione, perchè si occupa della gestione e manutenzione del datawarehouse e dell'elaborazione delle informazioni attraverso analisi di vario tipo ma ha un'importanza strategica: infatti dalle capacità dell'esperto di relazionarsi con i due mondi correttamente ed efficacemente dipende il successo dell'ambiente.



## APPENDICE A

### Il modello CRISP-DM

Il *Cross Industry Standard Process* per il *Data Mining* è un modello per la standardizzazione del processo di sviluppo e applicazione del data mining nei contesti organizzativi di tipo industriale e commerciale sviluppato nella metà degli anni '90 da un gruppo di esperti di data mining.

Il modello è nato da un workshop in cui si sono incontrati esperti di data mining per condividere gli approcci metodologici che ognuno aveva sviluppato individualmente nelle proprie esperienze lavorative e per fare il punto della situazione sullo stato del data mining applicato a qualsiasi contesto organizzativo; la domande che tali esperti si sono posti sono:

- *è possibile sviluppare un modello standard di processo di data mining che serva da guida a tutti i nuovi sviluppatori di processi di data mining?*
- *Come si dimostra ai nostri clienti che il data mining è giunto ad una maturità tale che è possibile adottarlo come parte di un loro processo di business chiave?*

La risposta alle due domande è lo sviluppo di uno modello di processo standard di data mining applicabile ad ogni contesto organizzativo e in ogni settore industriale; il modello è di riferimento per gli ideatori perché unisce le esperienze di ognuno e le standardizza in un una procedura unica ed è di riferimento per tutti i futuri sviluppatori sia di progetti grandi e complessi sia per quelli più piccoli. Le specifiche del modello sono disponibili a tutti e non proprietarie. I redattori del modello CRISP-DM incontratisi nel workshop hanno formato un consorzio che continua ancora oggi ad esistere e ad essere un punto di riferimento per tutti coloro che gravitano nel mondo del data mining.

Il modello di processo per il data mining dà una visione d'insieme della forma ciclica di un progetto di data mining, delle fasi in cui si articola il progetto, delle attività di ogni

fase e dei collegamenti tra le varie fasi. Il modello consiste di sei fasi, le fasi sono ordinate ma ogni fase può essere ripetuta all'interno del processo a seconda del contesto in cui si sta applicando il processo, il processo, nella sua interezza, è incrementale. I collegamenti tra le fasi vengono mostrati in figura 11.

Il modello fornisce le indicazioni delle direzioni, gli obiettivi e le attività di ogni fase del processo.

Il processo illustrato è di tipo top-down e gerarchico, il modello parte cioè da un livello di astrazione alto e quindi indipendente rispetto al contesto specifico e mano a mano che si scende nella gerarchia si trovano i riferimenti su che cosa fare e sul come farlo nei contesti specifici. Ogni fase è composta di un certo numero di obiettivi generici applicabili indipendentemente dal contesto del progetto, per ogni obiettivo sono definite una serie di attività da eseguire nella situazione specifica del processo, infine, all'ultimo livello di dettaglio, il processo CRISP è composto dalle varie istanze costituite dalle azioni, decisioni e risultati del processo di data mining studiato.

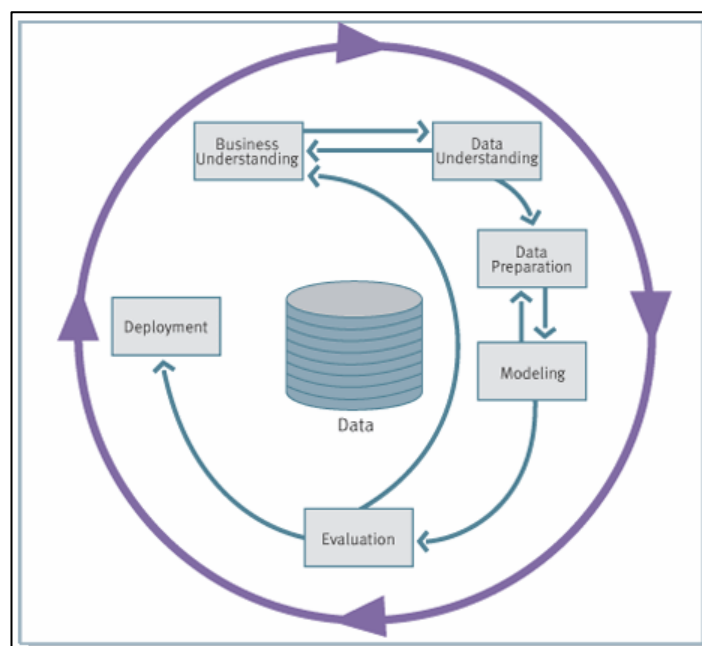


Figura 1: Fasi del modello CRISP-DM [fonte [www.crisp-dm.org](http://www.crisp-dm.org)]

## **Fase 1: *Business Understanding***

Questa fase iniziale consiste nel riflettere e considerare gli obiettivi del progetto e i suoi requisiti dal punto di vista aziendale.

La conoscenza accumulata in questa fase è necessaria per capire che cosa vuole veramente l'azienda dal data mining e se quello che vuole è compatibile con i vincoli che l'azienda stessa pone. Una volta chiariti gli obiettivi del business li si deve convertire in problemi di data mining; deve cioè essere definito che cosa ci si aspetta come output del sistema. Ad esempio, se l'obiettivo di business è di aumentare le vendite di un certo prodotto, il corrispondente obiettivo di data mining sarà di predire quante persone contattare nell'offerta e con quali caratteristiche, considerando i loro panieri di prodotti acquistati nel passato, le loro caratteristiche socio-demografiche e il loro stadio di relazione con l'azienda.

Infine si deve preparare una pianificazione molto dettagliata, che potrà cambiare nel corso del processo, delle attività da svolgere, delle risorse coinvolte e dei risultati attesi.

La fase 1 si compone degli obiettivi di:

1. determinare gli obiettivi di business;
2. conoscere la situazione aziendale in cui si stabilisce il processo;
3. determinare gli obiettivi di data mining;
4. produrre un piano ben definito del progetto.

## **Fase 2: *Data Understanding***

Questa fase inizia con la collezione dei dati provenienti dalle varie fonti e continua con tutte le attività di osservazione necessarie per prendere familiarità con i dati. Quando i dati provengono da fonti diverse è necessario verificare che sarà possibile integrarli, è necessario anche, indipendentemente dalla fonte, controllare la qualità dei dati, la sua numerosità e correttezza.

Questa fase è molto importante perché apporta una grande quantità di pre-conoscenza a partire dai dati e individua i possibili vincoli nell'applicazione del data mining per



soddisfar gli obiettivi prefissati nella fase precedente di *business understanding* dovuti alle caratteristiche dei dati.

La qualità del lavoro compiuto per la comprensione dei dati rappresenta il primo passo della fase successiva di preparazione dei dati.

Le attività di questa fase sono:

1. collezionare i dati iniziali;
2. descrivere i dati;
3. esplorarli;
4. verificarne la qualità.

### **Fase 3: *Data Preparation***

Questa fase comprende tutte le attività necessarie a costruire l'insieme dei dati che verrà sottoposto allo strumento di data mining, è pertanto necessario che a questo punto si sia arrivati all'idea definitiva di quali e quanti dati dovranno essere sottoposti all'analisi. Questa fase è solitamente molto lunga, può diventare automatizzabile una volta deciso che cosa si vuole dai dati e non presenta un ordine preciso di priorità tra le vari attività che la compongono:

1. selezione dei dati;
2. pulizia dei dati;
3. costruzione dei dati (record generati, variabili derivate);
4. integrazione dei dati;
5. formattazione dei dati in base al formato richiesto dallo strumento di data mining.

### **Fase 4: *Modeling***

In questa fase vengono selezionate le tecniche di modellazione ed applicate sui dati nella forma richiesta; nei risultati della modellazione hanno grande rilevanza le scelte degli

algoritmi ma soprattutto le scelte dei parametri che calibrano l'algoritmo. Di solito ci sono varie tecniche per lo stesso problema di data mining ed ogni tecnica presenta dei vincoli diversi sulla tipologia dei dati che prende in input. Dopo la scelta della tecnica di data mining da applicare può essere necessario pertanto dover tornare alla fase precedente di preparazione dei dati.

Questa fase comprende anche la valutazione del modello da un punto di vista tecnico di accuratezza e qualità dei risultati ottenuti; si valuta la bontà del modello sotto punti di vista diversi, non solo l'accuratezza con cui il modello predice un fatto ma anche la generalità dei risultati ottenuti, la comprensibilità e la semplicità del modello.

Le attività di questa fase sono:

1. selezione della tecnica di modellazione;
2. generazione della procedura o del meccanismo con cui testare il modello;
3. costruzione del modello;
4. miglioramenti al modello (valutazione tecnica del modello).

### **Fase 5: *Evaluation***

La fase di valutazione è strettamente legata agli obiettivi stabiliti nella fase di *business understanding*, se si crede che tali obiettivi non siano stati pienamente raggiunti è necessario ritornare sui passi precedenti.

Le attività di cui consiste questa fase sono:

1. valutazione di risultati;
2. review del processo;
3. definizione dei prossimi passi.

### **Fase 6: *Deployment***

L'ultimo passo del processo è quello di trasferire la conoscenza emersa dai modelli costruiti all'organizzazione. La conoscenza deve essere organizzata e presentata in maniera tale che sia comprensibile ed utile all'azienda.

Le attività di questa fase sono:

1. pianificare il trasferimento della conoscenza;
2. pianificare le azioni di monitoraggio e manutenzione del modello costruito;
3. produzione di un report finale;
4. ripercorrere ed eventualmente modificare tutte le fasi del processo.

## Bibliografia

M. Berry, G. Linoff [1997]  
*Data mining techniques: for marketing, sales, and customer relationship management*  
Wiley pub.

M. Berry, G. Linoff [2000]  
*Mastering data mining*  
Wiley pub.

M. Costabile [2001]  
*Il capitale relazionale*  
Mc Graw-Hill ed.

R. Daft [2001]  
*Organizzazione aziendale*  
Apogeo ed.

N. Del Ciello, S. Dulli, A. Saccaridi [2000]  
*Metodi di data mining per il customer relationship management*  
F. Angeli ed.

R. Drozdenko, P. Drake [2002]  
*Optimal database marketing*  
Sage pub.

P. Greenberg [2001]  
*CRM: customer relationship management*  
Apogeo ed.

C. Humby, T. Hunt, T. Phillips [2003]  
*Scoring points: how Tesco is winning customer loyalty*  
Kogan Page ed.

R. Kimball, M. Ross [2003]  
*Datawarehouse. La guida completa*  
Hoepli ed.

L. Milind [1993]  
*Il cliente è la chiave*  
F. Angeli ed.

E. Valdani [2000]  
*L'impresa pro-attiva*  
McGraw-Hill ed.

## **Bibliografia in rete**

[www.businessobjects.com](http://www.businessobjects.com)

[www.coop.it](http://www.coop.it)

[www.crisp-dm.org](http://www.crisp-dm.org)

[www.sas.com](http://www.sas.com)

[www.spss.com](http://www.spss.com)

[www.zara.com](http://www.zara.com)

[www.zegna.com](http://www.zegna.com)