

# Business Intelligence

## Il data mining in



L'analisi matematica per dedurre schemi e tendenze dai dati storici esistenti.

- **Revenue Management. Previsioni di occupazione.**
- **Marketing. Mail diretto a clienti specifici.**

**Gilda Camarovi**  
Developer Evangelist  
Developer & Platform Evangelism  
**Proxima Service srl**

## Introduzione

**HotelCube International** è il programma di gestione alberghiera (PMS) più innovativo sul mercato sia dal punto di vista tecnologico che funzionale. Infatti, oltre a tutte le necessarie funzioni presenti nei software per le strutture ricettive, le funzioni di **Rate Management**, **Revenue Management**, **CRM**, **Marketing**, l'integrazione con **Booking Engine** per l'e-commerce e la gestione centralizzata di tutti documenti (corrispondenza, conferme prenotazioni, pratiche...), tramite l'integrazione con Outlook 2007 e Word 2007, rendono **HotelCube International** lo strumento indispensabile per la gestione manageriale del Vostro albergo.

Il presente documento è rivolto a tutti coloro che sono interessati a conoscere le nuove funzionalità di **HotelCube International** e l'apprendimento dei risultati forniti dalle funzioni di:

- ◆ **Previsione** del modulo di **Revenue Management**
- ◆ **Marketing** del modulo **CRM (Customer Relationship Management)**

Pur utilizzando modelli matematici e funzionalità molto complesse, **HotelCube International** non richiede la conoscenza tecnica di tali strumenti. Il Marketing Manager, il Revenue Manager ed il Direttore troveranno valide indicazioni per poter valutare la potenzialità degli strumenti e l'aderenza degli stessi alle crescenti necessità del management delle strutture ricettive.

## Contenuto del documento

La grande disponibilità di dati, provenienti dal gestionale (PMS), dal web, dal CRM o da altre sorgenti e le tecniche di analisi permette una migliore comprensione delle tendenze dei clienti ed un utilizzo più semplice dei risultati nei processi decisionali. Da tempo, infatti, gli "statistici" si sono occupati di studiare e fornire strumenti matematici per "stimare" i comportamenti di un "collettivo". Gli schemi e le tendenze possono essere raccolti e definiti in un *modello di data mining*.

**Il data mining, in sintesi, ha per oggetto l'estrazione di dati (attraverso metodi automatici) e l'utilizzazione operativa di questo "sapere".**

**HotelCube International** elabora i dati del PMS utilizzando specifici modelli di **data mining**, realizzati da Microsoft e codificati nel motore del database, ed i risultati delle elaborazioni vengono utilizzati dalle funzionalità di CRM e Revenue Management.

Questo documento è suddiviso in due parti e si propone di presentare due modelli matematici utilizzando due case history" in cui l'applicazione delle funzionalità del **data mining** di **HotelCube International** riveste un'importanza fondamentale per ottenere indicazioni indispensabili per la gestione manageriale.

- ◆ **Parte I. Previsioni di Occupazione** - Lo strumento per il Revenue Manager.
- ◆ **Parte II. Mailing diretti a clienti specifici** - Il giusto messaggio promozionale.

Il documento è solo una presentazione delle funzionalità e non vuole essere una trattazione esaustiva delle funzionalità, dei risultati o delle tecniche utilizzate.

## Parte I. Previsioni di Occupazione

Per la corretta massimizzazione del Revenue il modello più usato è quello del *Bid Price*. In tale contesto la massimizzazione del ricavo (la nostra *funzione obiettivo*) in un preciso giorno comporta la determinazione di un numero preciso di camere da riservare a ciascun Segmento di Business (e quindi Classe Tariffaria) tenendo in considerazione la capacità ricettiva globale dell'albergo e la previsione di domanda. E' necessario, quindi, per una corretta applicazione del **Revenue Management** prevedere l'occupazione dell'albergo (breve periodo e lungo periodo). Per poter prevedere (stimare, come è preferibile affermare) è necessario disporre della *serie storica* dell'occupazione, ovvero l'occupazione che si è avuta giorno dopo giorno. Per la stima dell'occupazione prevista **HotelCube International** si basa sul modello stocastico di Box e Jenkins (comunemente chiamato ARIMA che è l'acronimo di Auto Regressive Integrated Moving Average). Il modello viene applicato in due fasi automatiche. Nella prima "apprende" l'andamento dell'occupazione in funzione dei dati storici, nella seconda fornisce la stima. Il ciclo può essere ripetuto più volte, apportando i "correttivi" necessari (eventi che hanno portato ad un'occupazione eccezionalmente alta o bassa ma che sono frutto di cause impreviste ed imprevedibili. Per esempio una partita di Champions League o uno sciopero del trasporto aereo).

### Case history

Albergo di 150 camere che utilizza in modo appropriato la Segmentazione del Cliente. La Segmentazione è un elemento indispensabile per il Revenue Management. Infatti lo stesso prodotto (camera e servizi accessori) sarà venduto a tariffe diverse in funzione della capacità di spesa del Cliente. I segmenti vanno controllati periodicamente per verificare se il comportamento dei clienti nel tempo è cambiato. **HotelCube International** mette a disposizione modelli personalizzabili per la verifica dei segmenti (profili clienti), ma questo tipo di funzionalità esula dall'argomento trattato in questo paragrafo. Il Revenue Manager esegue, giorno per giorno, la stima dell'occupazione e formula l'ottimizzazione del numero di classi tariffarie disponibili per la vendita.

Le figure e le tabelle riportate nelle pagine successive sono solo una piccolissima parte di quelle fornite da **HotelCube International**. Questa collezione di output permette al Revenue Manager di valutare i risultati e prendere le decisioni migliori.

La Fig. I.1 visualizza l'andamento dell'occupazione dell'albergo negli anni 2005, 2006, 2007 e 2008 ottenuta con la funzionalità di "Analisi Serie Storiche" di **HotelCube International**. E' possibile, con un semplice click del mouse, eseguire l'azione di "drill through", ovvero, un'analisi in dettaglio a partire da un dato aggregato. La Fig. I.2 riporta il dettaglio dell'occupazione del mese di marzo 2008 per Segmento di Cliente, la Fig. I.3 riporta il dettaglio per ciascun giorno del mese di marzo del 2008 sempre per Segmento Cliente.

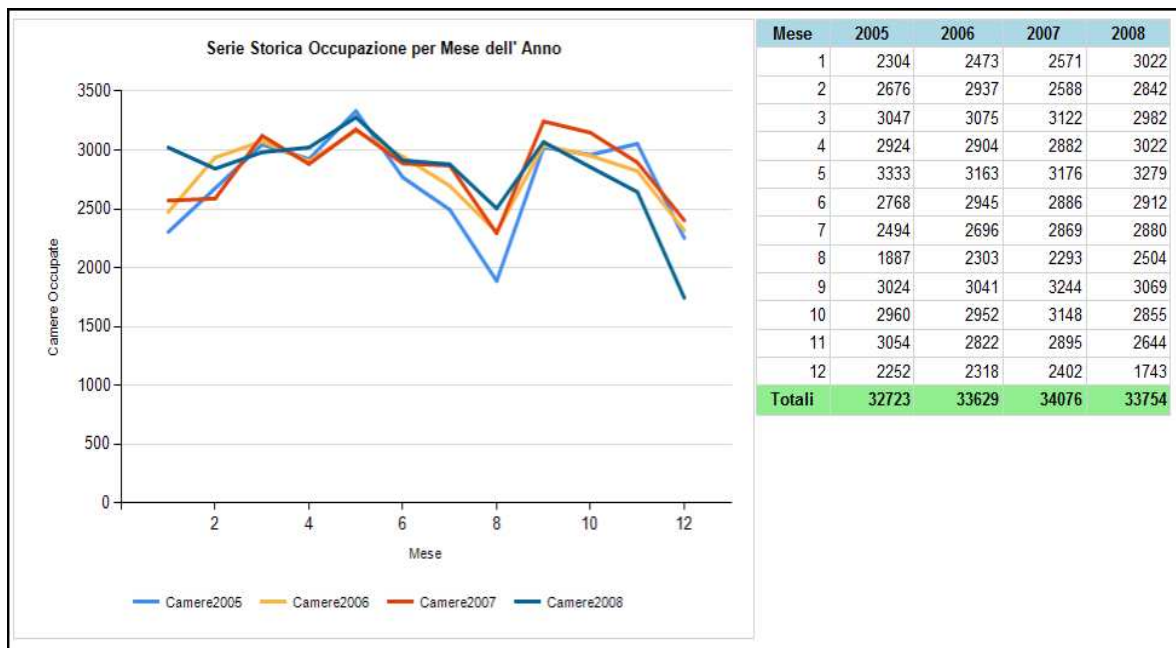


Fig. I.1

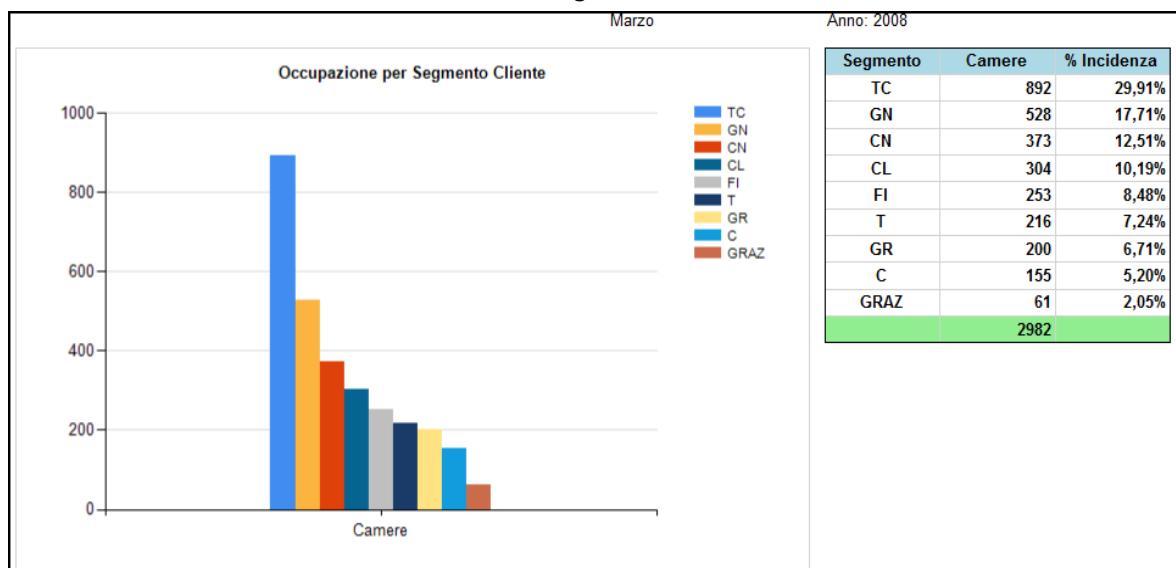


Fig. I.2

Segmento	01/03/2008	02/03/2008	03/03/2008	04/03/2008	05/03/2008	06/03/2008	07/03/2008	08/03/2008	09/03/2008
TC	29	24	22	36	36	40	34	42	38
FI	17	2	1					1	2
GN	9					26	17	26	25
GR	6		9	14	14		7	7	
CN	6	7	23	28	26	23	15	14	6
CL	5	4	19	12	9	5	1	2	4
T	5	8	3	1	1	1	15	15	10
GRAZ		7	11	11	6	18	7		
C		7	17	18	13	7	1		
	77	59	105	120	105	120	97	107	85

Fig. I.3

L'analisi dell'andamento dei dati evidenzia l'esistenza di una componente di lungo periodo (*trend*) e di una componente stagionale che rende la serie *non stazionaria in media e in varianza*. Prima di far eseguire al risolutore del modello matematico la fase di "apprendimento" e successivamente la previsione, è necessario stabilizzare la variabilità, che è un prerequisito del modello ARIMA. Queste operazioni sono totalmente automatizzate e, alla fine dell'elaborazione, si otterrà la stima per ogni mese del 2009 (Fig. I.4). Nella stessa figura viene visualizzata l'occupazione del 2008 e l'occupazione effettiva registrata fino al 15 marzo 2008. La Fig. I.5 visualizza graficamente l'andamento dell'occupazione e la relativa stima (la curva tratteggiata).

Utilizzando la stessa funzione per la soluzione del modello è possibile ottenere la stima dell'occupazione giornaliera del mese di Marzo 2009. La Fig. I.6 visualizza l'occupazione registrata nel mese di Marzo degli anni 2005, 2006, 2007 e 2008. La Fig. I.7 visualizza graficamente l'andamento dell'occupazione e la previsione (la linea tratteggiata). La Fig. I.8 visualizza giorno per giorno la previsione del mese di Marzo 2009, l'occupazione registrata del mese di Marzo 2008 e l'occupazione registrata fino al 15 Marzo 2009. L'ulteriore e più significativo risultato è quello della stima per Segmento Cliente (e quindi Classe Tariffaria). La Fig. I.9 visualizza la stima ottenuta per ciascun giorno del mese di Marzo 2009 per ciascun Segmento Cliente (C=Ditte convenzionate, GR=Gruppo, T=Turistico ...). La Fig. I.10, infine, visualizza l'andamento storico e di previsione relativa al segmento C.

Al Revenue Manager non sfugge l'importanza della segmentazione dell'offerta. In **HotelCube International** sono disponibili specifiche funzioni per la determinazione dei profili dei comportamenti dei clienti (capacità di spesa, lead time, long stay, ecc...) e relativa segmentazione.

**HotelCube International** fornisce tutte le funzionalità per applicare il **Revenue Management** nel vostro Albergo (Segmentazione, Previsione, Ottimizzazione, Vendita e Controllo).

Previsione Occupazione per l' anno					
2008	Storico HCI	2009	Previsione HCI	2009	Attuale HCI
01/2008	3022	01/2009	2872	1	2159
02/2008	2842	02/2009	2901	2	2456
03/2008	2982	03/2009	3039	3	1062
04/2008	3022	04/2009	2722	<b>Totale</b>	<b>5677</b>
05/2008	3279	05/2009	2647		
06/2008	2912	06/2009	2688		
07/2008	2880	07/2009	2876		
08/2008	2504	08/2009	2885		
09/2008	3069	09/2009	2914		
10/2008	2855	10/2009	2774		
11/2008	2644	11/2009	2689		
12/2008	1743	12/2009	2622		
<b>Totale</b>	<b>33754</b>	<b>Totale</b>	<b>33629</b>		

Fig. I.4

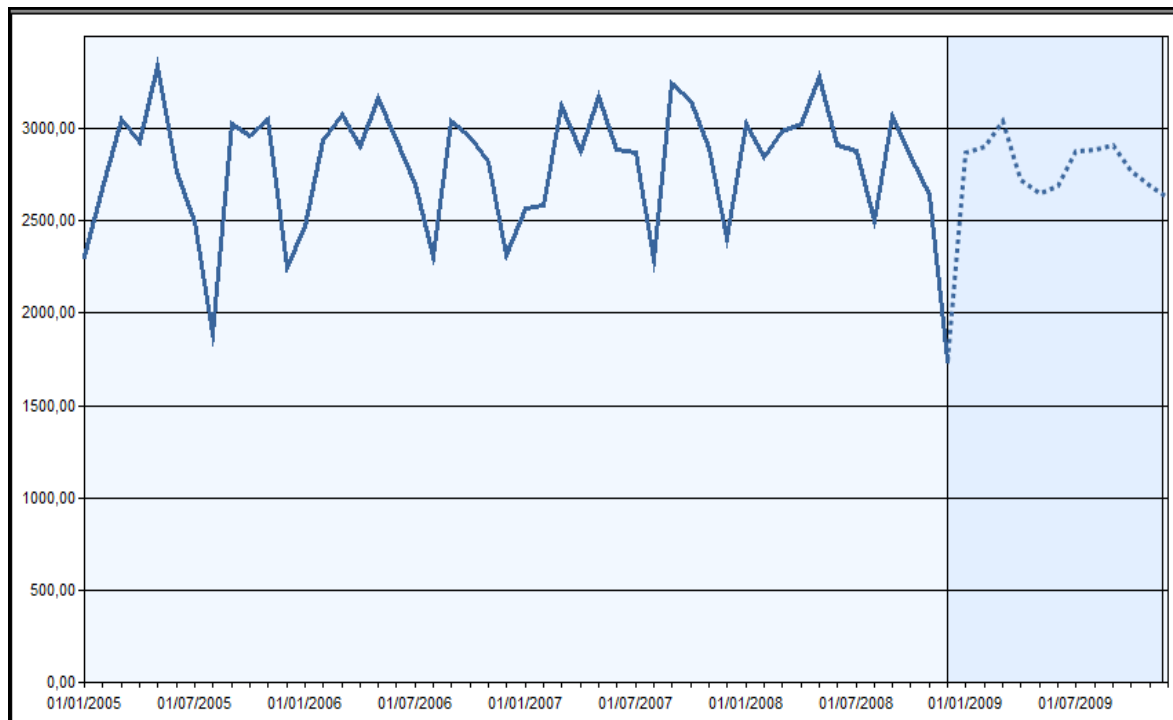


Fig. I.5

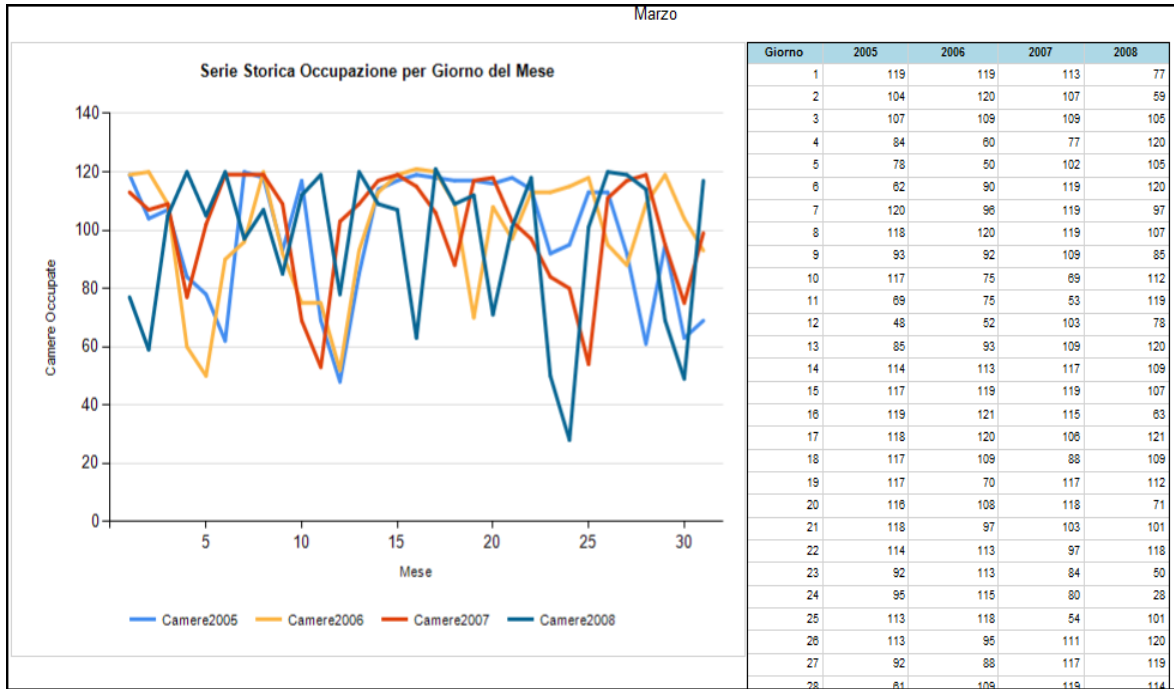


Fig. I.6

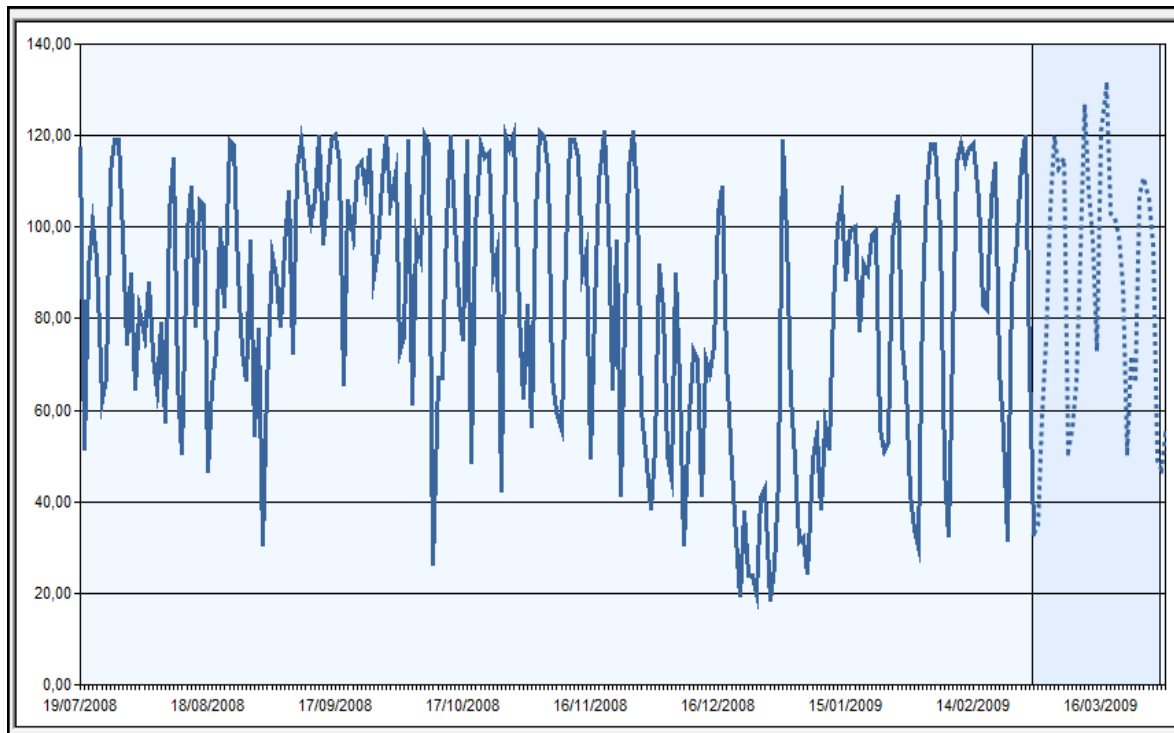


Fig. I.7



Previsione Occupazione mese di : Marzo					
2008	Storico HCI	2009	Previsioni HCI	2009	Attuale HCI
01/03/2008	77	01/03/2009	35	01/03/2009	37
02/03/2008	59	02/03/2009	59	02/03/2009	49
03/03/2008	105	03/03/2009	76	03/03/2009	94
04/03/2008	120	04/03/2009	104	04/03/2009	119
05/03/2008	105	05/03/2009	120	05/03/2009	119
06/03/2008	120	06/03/2009	113	06/03/2009	90
07/03/2008	97	07/03/2009	115	07/03/2009	83
08/03/2008	107	08/03/2009	51	08/03/2009	57
09/03/2008	85	09/03/2009	56	09/03/2009	58
10/03/2008	112	10/03/2009	66	10/03/2009	72
11/03/2008	119	11/03/2009	95	11/03/2009	80
12/03/2008	78	12/03/2009	127	12/03/2009	71
13/03/2008	120	13/03/2009	108	13/03/2009	31
14/03/2008	109	14/03/2009	97	14/03/2009	73
15/03/2008	107	15/03/2009	73	15/03/2009	29
16/03/2008	63	16/03/2009	121	<b>Totale</b>	<b>1062</b>
17/03/2008	121	17/03/2009	131		
18/03/2008	109	18/03/2009	103		
19/03/2008	112	19/03/2009	102		
20/03/2008	71	20/03/2009	98		
21/03/2008	101	21/03/2009	87		
22/03/2008	118	22/03/2009	51		
23/03/2008	50	23/03/2009	72		
24/03/2008	28	24/03/2009	66		
25/03/2008	101	25/03/2009	107		

Fig. I.8

Previsione Occupazione segmenti mese di : Marzo									
Segmento	04/03/2009	05/03/2009	06/03/2009	07/03/2009	08/03/2009	09/03/2009	10/03/2009	11/03/2009	12/03/2009
C	14	18	12	6	17	20	12	7	
CL	25	18	1	4	6	24	25	23	
CN	14	14	7	3	3	11	15	15	
FI	18	12	7	1	4	16	16	16	
GN	12	18	20	21	18	9	7	8	
GR	19	19	23	24	25	6	18	14	
GRAZ	5	4	4	4	3	3	3	3	
MTGO	1	2	1	2	0	2	0	2	
T	7	10	10	6	4	10	3	6	
TC	9	9	15	21	19	12	10	18	
	<b>104</b>	<b>120</b>	<b>113</b>	<b>115</b>	<b>51</b>	<b>56</b>	<b>66</b>	<b>95</b>	

Fig. I.9

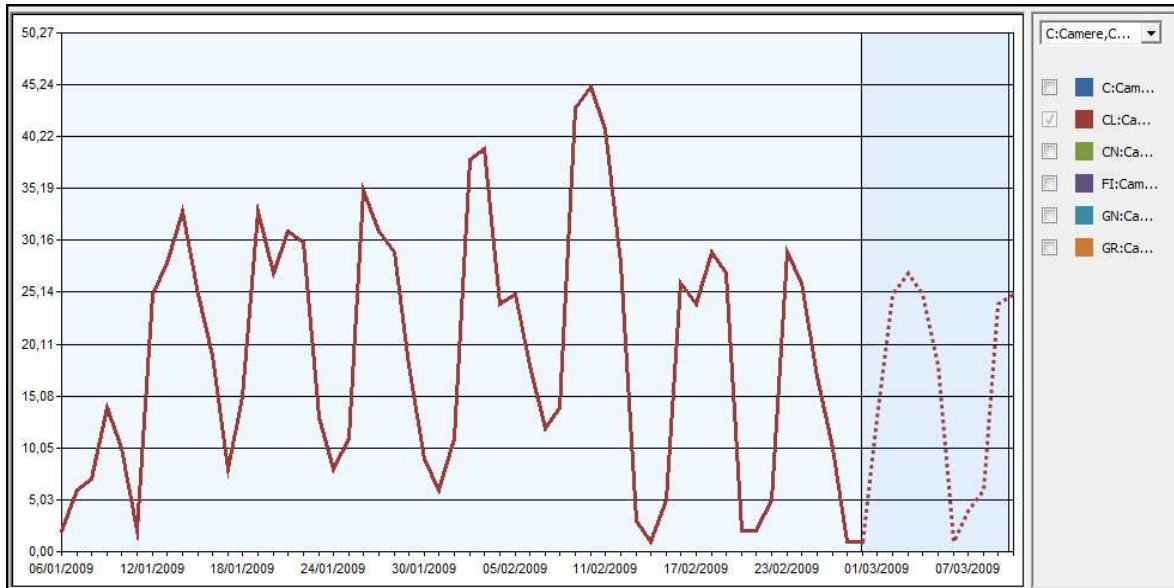


Fig. I.10

Le previsioni così ottenute vengono registrate automaticamente nel database per essere utilizzate nella fase di Ottimizzazione. Si desidera ribadire che la previsione è un processo di stima che non può tenere conto dei fattori (e sono tantissimi ed aleatori) che possono incidere tanto positivamente quanto negativamente sull'occupazione. Per tale motivo nella funzione di ottimizzazione è possibile modificare i dati stimati sulla base delle conoscenze, dell'intuito e delle analisi "a vista" del Revenue Manager.

## Conclusioni

La funzionalità di previsione e la collezione di statistiche ed analisi (presentate in questo documento solo in minima parte) fanno parte integrante di **HotelCube International**.

I nostri trainer Vi forniranno tutte le informazioni per il miglior utilizzo.

## Parte II. Mailing diretto a clienti specifici

Il Vostro reparto Marketing intende aumentare le vendite inviando messaggi promozionali a specifici clienti. Il database dell'azienda contiene un elenco dei clienti acquisiti ed un elenco di potenziali nuovi clienti registrati tramite la gestione delle *attività di marketing* del modulo **CRM** di **HotelCube International**. In base all'esame delle tipologie dei precedenti acquirenti di prestazioni alberghiere e SPA, l'Albergo vuole individuare modelli che siano applicabili alla clientela potenziale. L'obiettivo è quello di individuare modelli che consentano di eseguire previsioni circa i potenziali clienti che con maggiore probabilità acquisteranno un pacchetto o una prestazione del Vostro Albergo.

### Case history

Albergo che raccoglie ed inserisce dati aggiuntivi rispetto a quelli classici rilevabili dal check in. Tali dati sono utili sia per la fidelizzazione del cliente (memorizzare se l'ospite è un fumatore o no, se richiede un posto specifico in garage perché ha una auto voluminosa, il giornale che desidera leggere al mattino, se è vegetariano, ecc... può rendere la fase di booking più accattivante, portando alla constatazione da parte dell'ospite di essere "riconosciuto") sia per le statistiche utilizzate dai modelli di marketing. Il reparto Marketing, inoltre, raccoglie una serie di dati sui potenziali clienti ed ospiti, memorizzandoli tramite le funzioni di "Gestione Contatti ed Attività". Per questo esempio di analisi, l'Albergo intende utilizzare i dati suddivisi nei tre gruppi elencati di seguito.

Dati Anagrafici	Dati Presenze	Dati aggiuntivi
Cognome	N. volte ospitato	Professione
Nome	Capacità di spesa	Stato Civile
Città		
Indirizzo		
Cap		
Provincia		
Sesso		
Indirizzo e-mail		
Età		

L'algoritmo utilizzato è quello del Decision Trees e le tabelle saranno quella degli ospiti (che rappresenta la tabella dei *cas*) e la tabella dei contatti.

La prima cosa da definire è quale informazione si vuole considerare, "stimabile" e di "input", tra quelle utilizzate dal modello nella fase di apprendimento dei dati.

La Fig. II.1 visualizza la struttura del modello.

Struttura modello di data mining:

Table/Colonne	Chiave	Input	Stimabile
vClientiMKTG			
CapResidenza	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CittaResidenza	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CodiceAnagrafica	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cognome	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
DataNascita	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eta	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
IndirizzoEMail	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
IndirizzoResidenza	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nome	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NVolte	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Professione	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
ProvinciaResidenza	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Sesso	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
StatoCivile	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Fig. II.1

L'algoritmo utilizzato permette di stimare quali "colonne" (ovvero i campi che abbiamo incluso) influenzano sulla decisione di soggiornare in albergo.

Effettuata l'elaborazione, le funzioni che consentono di esplorare i risultati del modello sono:

- ◆ **Albero delle decisioni**
- ◆ **Rete di dipendenze**

La Fig. II.2 visualizza l'albero delle decisioni a tre livelli (il risultato dell'elaborazione indica la presenza di 48 livelli che possono essere visualizzati con un semplice click sul simbolo +) da cui è possibile dedurre che l'età è il fattore più importante nella stima dell'acquisto del soggiorno.

Un dato interessante da sottolineare è che, una volta raggruppati i clienti per età, il ramo successivo dell'albero è diverso per ogni nodo di età. Esplorando la scheda *Albero delle decisioni* è possibile concludere che le persone di età compresa tra 41 e 47 anni di sesso maschile acquisteranno un soggiorno con molta probabilità, allo stesso modo dei "single" di età più giovane che vivono nel nord-est.

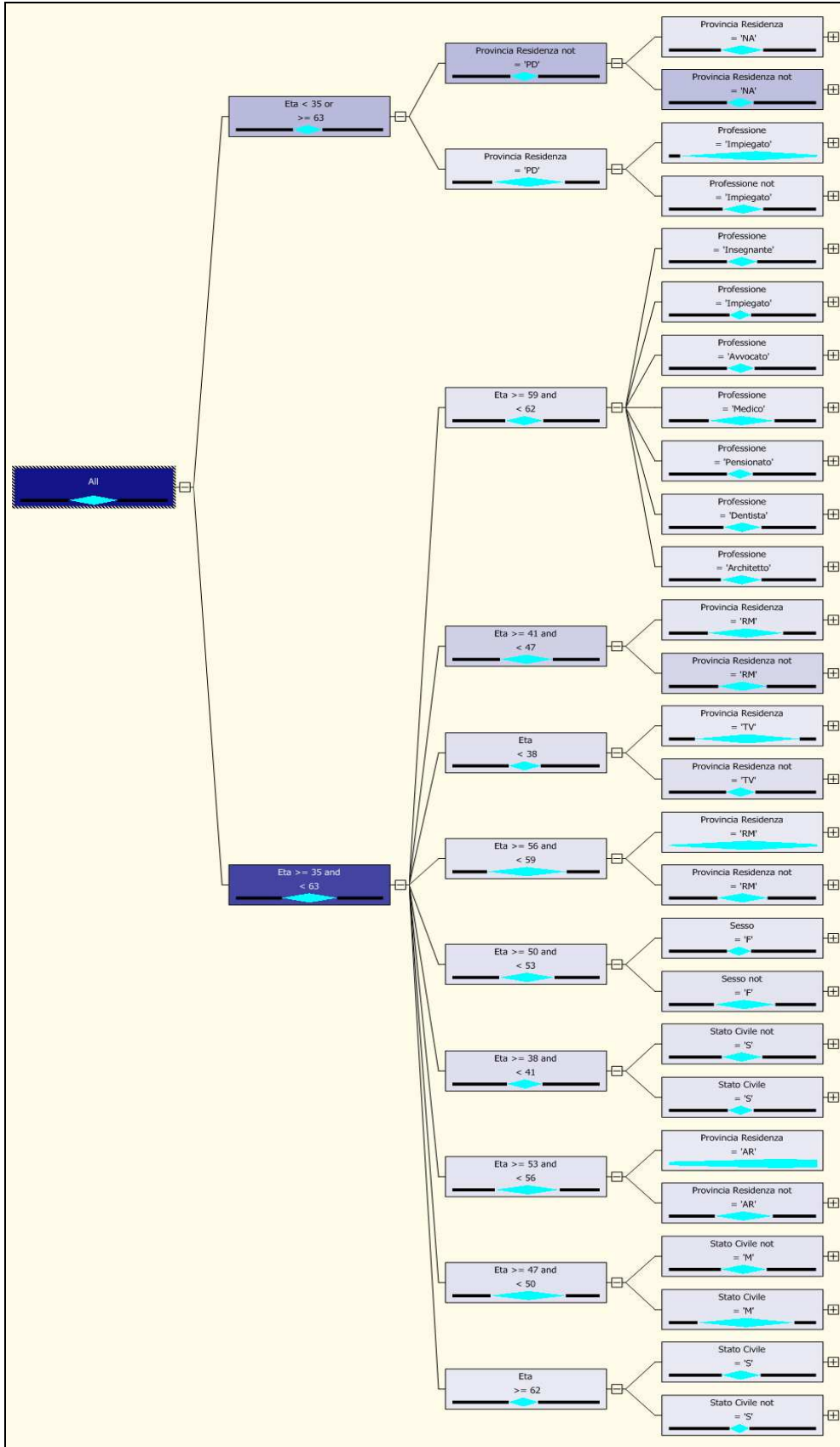


Fig.II.2

La figura sottostante (Fig. II.3) è un parziale zoom della visualizzazione precedente per fornire una visione più chiara dei dati rappresentati.

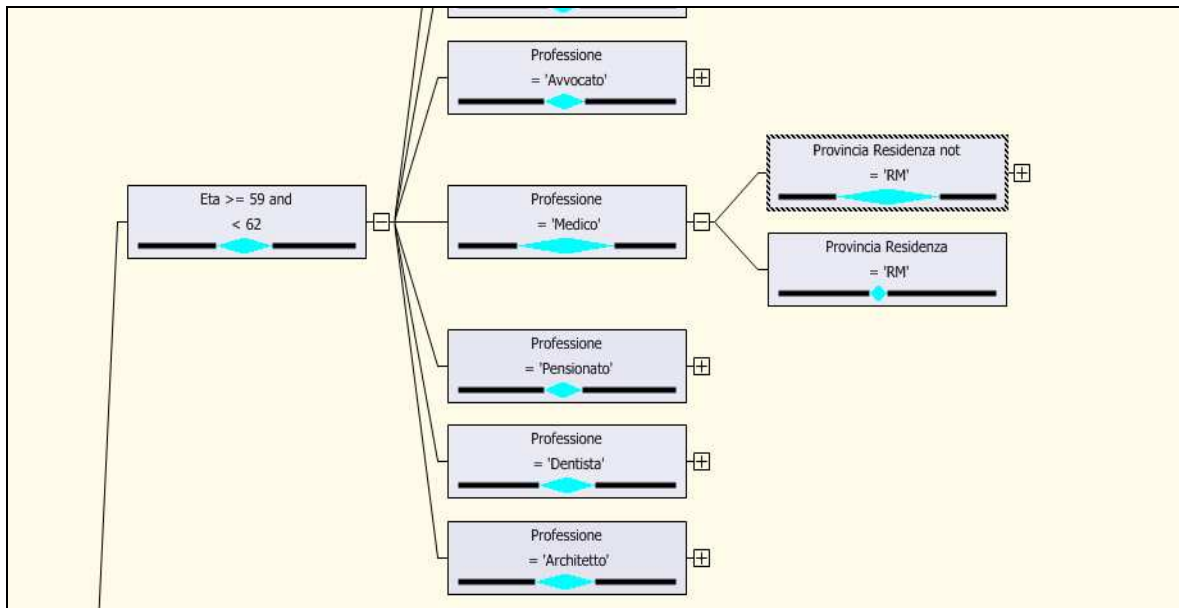


Fig. II.3

La Fig. II.4, infine, visualizza la *Rete delle dipendenze*. La funzione fornisce, attraverso un cursore, la visualizzazione dei collegamenti più attendibili. Spostando il cursore da poco attendibile (il caso rappresentato nella figura con tutti i collegamenti visibili che puntano sulla colonna stimabile) a più attendibile, le frecce che rappresentano relazioni meno attendibili diventeranno invisibili.

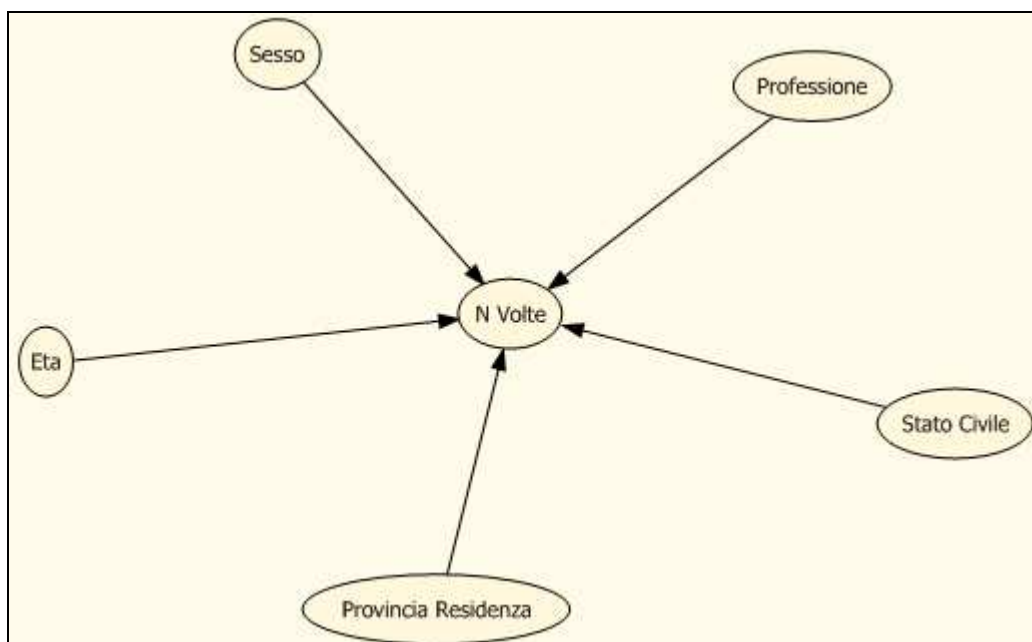


Fig. II.4

Il risultato finale dello spostamento del cursore che misura l'attendibilità delle relazioni è visualizzato nella Fig. II.4

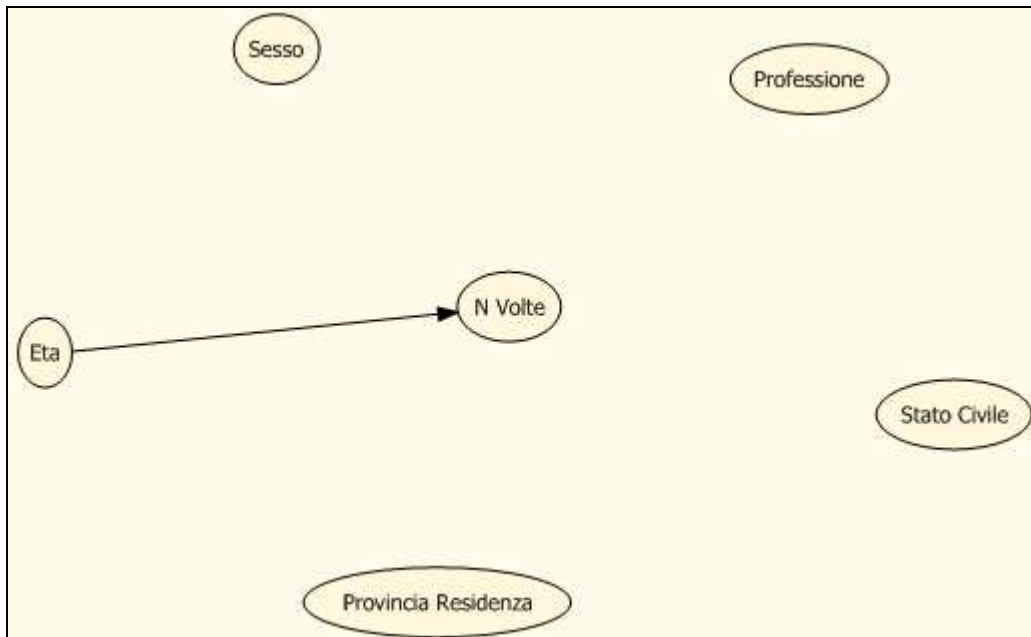


Fig. II.4

La sequenza dell'attendibilità delle relazioni ha eliminato, nell'ordine, la Professione, il Sesso, lo Stato Civile e la Provincia di Residenza.

## Conclusioni

Questa funzionalità è parte integrante di **HotelCube International**. I nostri trainer Vi forniranno tutte le informazioni per il miglior utilizzo.