

UNIVERSITA' DEGLI STUDI DI
PADOVA
FACOLTA' DI SCIENZE STATISTICHE

Corso di Laurea in
Statistica e Tecnologie Informatiche

**“Impiego di un software di Business Intelligence:
Il caso Domnick Hunter Hiross Spa”**

**Tesi di laurea di:
Rostellato Maurizio**

**Relatore:
Prof.sa Dulli Susi**

Anno accademico 2004/2005

INDICE:

Introduzione

CAPITOLO I

PRESENTAZIONE DELLO STAGE

CAPITOLO II

DOMNICK HUNTER HIROSS

CAPITOLO III

IL PROGETTO "BUSINESS INTELLIGENCE"

CAPITOLO IV

IL PROCESSO DI DATA WAREHOUSING

- 4.1 Dati
- 4.2 Sistemi OLTP
- 4.3 Sistemi OLAP
- 4.4 Mappatura
- 4.5 Estrazione dati
- 4.6 Aggiornamento incrementale del data warehouse
 - 4.6.1 Metodi statici
 - 4.6.2 Metodi dinamici
- 4.7 Trasformazioni fisiche
- 4.8 Trasformazioni logiche

CAPITOLO V

PROGETTO "CUBO VENDITE"

- 5.1 I Fatti
- 5.2 Le Dimensioni
- 5.3 Gli Attributi
- 5.4 Gerarchie degli attributi
- 5.5 Analisi dei requisiti tecnici
 - 5.5.1 Dimensioni
 - 5.5.2 Naming degli oggetti del database
 - 5.5.3 Dettaglio informativo
- 5.6 Disegno logico della base dati
 - 5.6.1 Definizione delle dimensioni di analisi
 - 5.6.2 Layout multidimensionali

CONCLUSIONI

Bibliografia

INTRODUZIONE

L'economia va incontro a un continuo processo di "dematerializzazione" contraddistinto dal primato del know-how sul capitale, che porta con sé due principali conseguenze: una riduzione dei costi, che si traduce in benessere per la società nel suo complesso, e la globalizzazione del mercato, favorita dalla facilità con cui le informazioni, i prodotti e i servizi riescono a circolare rispetto ai beni fisici.

Il calo dei costi e la globalizzazione del mercato implica una spinta concorrenziale senza precedenti e una graduale diminuzione di importanza dei confini geografici.

Cambia la natura della competizione, non essendo più sufficiente l'efficienza se non accompagnata da capacità innovativa, qualità, servizio e flessibilità.

In un mercato dominato dalle caratteristiche dell'ipercompetizione, dunque, le imprese si affermano grazie all'originalità e alle capacità innovative: la differenziazione della propria offerta, la diversificazione tra i prodotti, la presenza in aree geografiche nuove e distanti. In passato la competizione tra aziende era basata in larga misura sui differenziali di costo, ma, successivamente, l'ingresso, sul mercato dei paesi in via di sviluppo con costi di produzione molto più bassi e una discreta capacità industriale hanno causato un progressivo spostamento dell'attenzione sui differenziali di competitività e sulla capacità delle imprese di migliorare i propri prodotti e servizi.

Per soddisfare il fabbisogno di conoscenza dell'impresa occorre realizzare un sistema di business intelligence che costituisca un

efficace risposta per gestire il cambiamento ed operare con successo nel contesto competitivo.

Il significato di business intelligence è legato alla parola "intelligence" che è l'abilità ad imparare, a capire una nuova situazione, la capacità ad utilizzare la ragione, l'abilità ad applicare la conoscenza.

In altre parole la BI è un insieme di concetti, metodi e processi per supportare e migliorare le decisioni aziendali usando i dati e le informazioni provenienti da sorgenti multiple a cui viene applicata l'esperienza per sviluppare un'accurata comprensione delle dinamiche del business.

Le aree alle quali è possibile applicare le tecnologie di BI rappresentano la quasi totalità dei settori aziendali, anche se storicamente il CRM ha visto il maggior numero di applicazioni sul campo. Tanto più le aziende producono informazioni, tanto più hanno bisogno di presentarle in maniera armonica per supportare i processi decisionali; la BI si rivolge principalmente al knowledge worker, fornendo gli strumenti necessari a prendere decisioni e risolvere problemi. E' necessario disporre di un funzionale sistema informativo che permetta di raccogliere i dati provenienti dalle fonti interne ed esterne allo scopo di trasformarli in informazioni utili per supportare i processi decisionali. Molte imprese, infatti, anche se carenti di stock di risorse umane, finanziarie e tecnologiche riescono a competere brillantemente sul mercato in quanto possiedono ottime capacità in termini di flessibilità, tempi di reazione e capacità di innovazione.

In uno scenario complesso e instabile come quello della nuova economia, infatti, uno degli elementi fondanti è la conoscenza, ciò spiega l'interesse sempre maggiore delle imprese nei confronti del knowledge management. L'applicabilità pratica di sistemi per la creazione e gestione della conoscenza è diventata concreta solo

recentemente con il diffondersi di sistemi di accesso all'informazione automatizzati. I mezzi tecnologici odierni consentono di gestire la conoscenza con basi di dati più flessibili e facilmente condivisibili, con enormi capacità di memorizzazione, con accesso multimodale e a costi molto bassi.

Il knowledge management, tuttavia, non riguarda solo un ambito tecnologico, ma incide, soprattutto, sugli aspetti organizzativi interni all'impresa.

Difatti, anche utilizzando le più sofisticate infrastrutture, se non si attivano i meccanismi organizzativi, i processi interni e, soprattutto, le motivazioni, la cultura e i valori delle persone che lavorano in azienda e accettano di condividere la conoscenza, risulterà difficile ottenere una sua efficace gestione.

Le soluzioni di business intelligence flessibili ed integrate rappresentano, dunque, uno strumento gestionale efficace nei contesti concorrenziali perché permettono di completare il processo di trasformazione dei dati in informazioni e di queste ultime in conoscenza, allo scopo di conseguire un vantaggio competitivo durevole. Molte aziende, tuttavia, sono spesso frenate nell'investire in progetti di business intelligence, in quanto la valutazione ex ante degli investimenti non è quantificabile con precisione.

CAPITOLO I

LO STAGE

L'esperienza di stage in Domnick Hunter Hiross, di complessive quattrocentosettantacinque ore ripartite su cinque mesi (ottobre 2004 – febbraio 2005), si è svolta nell'area controllo di gestione che, come da funzionigramma aziendale (figura1), insieme all'area amministrativa è alle dipendenze del Finance Manager che a sua volta dipendente direttamente dalla direzione Operation.

La mia attività principale è stata di supporto ai tre Controller dell'ufficio, impegnati, oltre che nelle attività di routine (budgeting, reporting giornaliero, chiusure mensili, ecc..), nello sviluppo del progetto "Business Intelligence" in stretta collaborazione con l'Information Technology, e società di consulenza esterna (Deloitte, Reply).

I compiti a me assegnati, inizialmente, riguardavano l'attività di reporting, giornaliero e mensile, necessaria al monitoraggio delle variazioni delle vendite (fatturato, ordini, portafoglio) e di magazzino (movimenti, giacenze), la compilazione dei report era eseguita manualmente riportando su griglie statiche di Excel (figura3) dati ottenuti da tabelle pivot riferite a basi dati estratte tramite query dal sistema informativo transazionale JDEdward's, aggiornate tramite macro.

Con il completamento del primo step del progetto B.I. (cubo VENDITE), l'attività di reporting manuale è andata scemando, nell'ambito delle vendite, lasciando spazio alla creazione di nuovi report con il nuovo strumento Everest di Outlook Soft (figura4), il

quale permette di avere in modo automatico estrazioni di dati aggiornati e compilazione di report, costruiti in base alle richieste delle varie aree aziendali, con un notevole risparmio in termini di tempo da parte dei controller e un immediata reperibilità delle informazioni da parte degli utenti.

Il controllo delle anagrafiche (CLIENTE, PRODOTTO), altra attività a cui ho partecipato, è stata fondamentale per il corretto funzionamento dell'applicazione di Business Intelligence.

FUNZIONIGRAMMA DIREZIONE OPERATIONS
14 DICEMBRE 2004

Firma GM/OPS: *Aldo Moro*

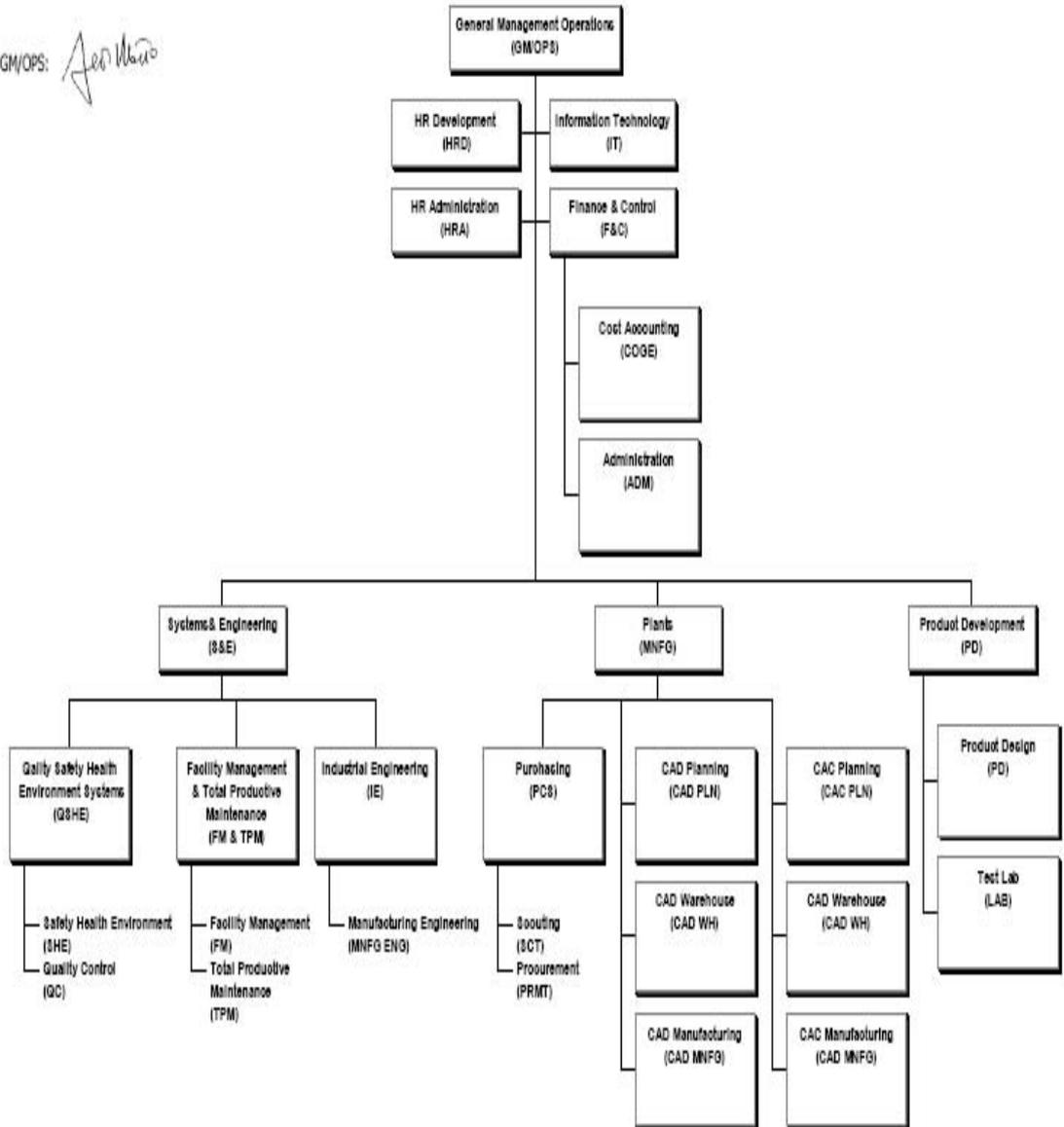


Figura 1: Funzionigramma DHH Operation

FUNZIONIGRAMMA SALES & MARKETING
14 DICEMBRE 2004

Firma GM/S&M: *Firma Fopola'*

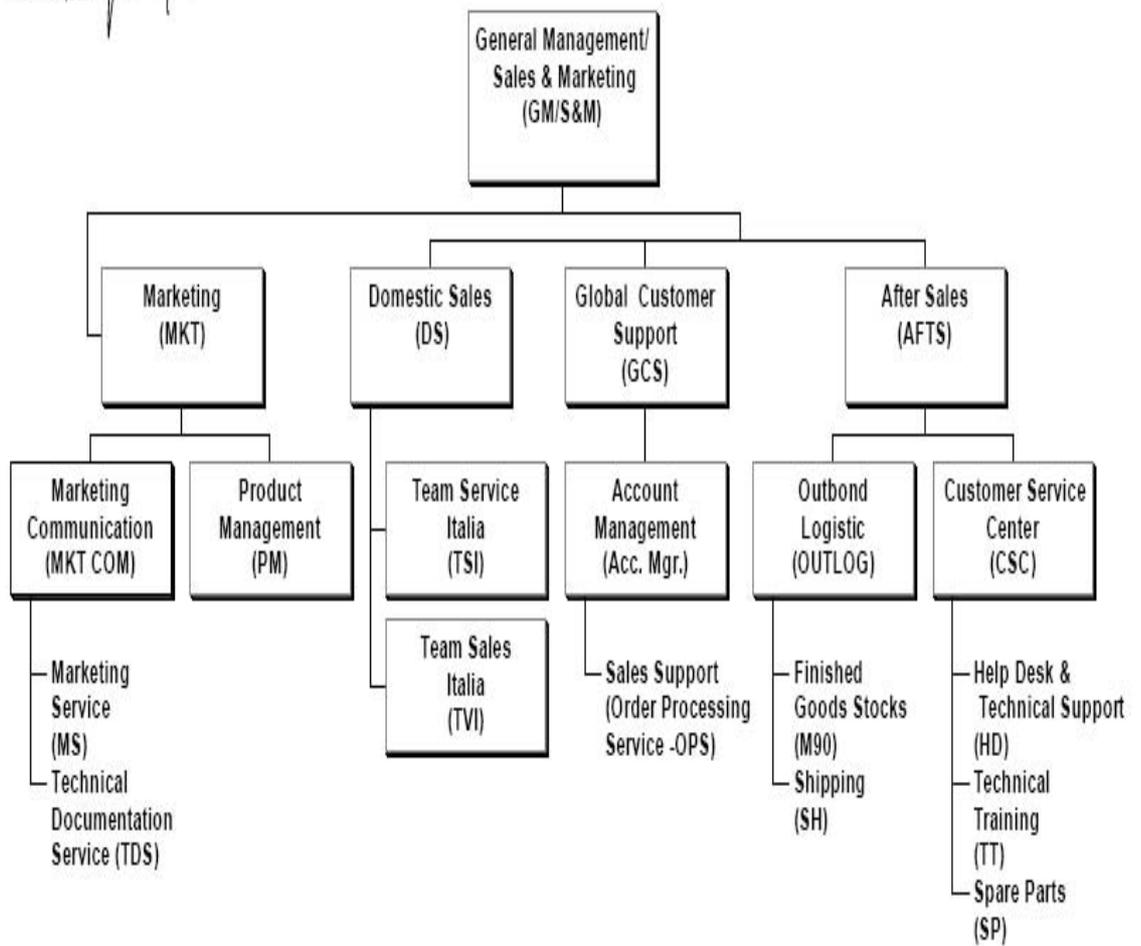


Figura 2: Funzionigramma DHH Sales & Marketing

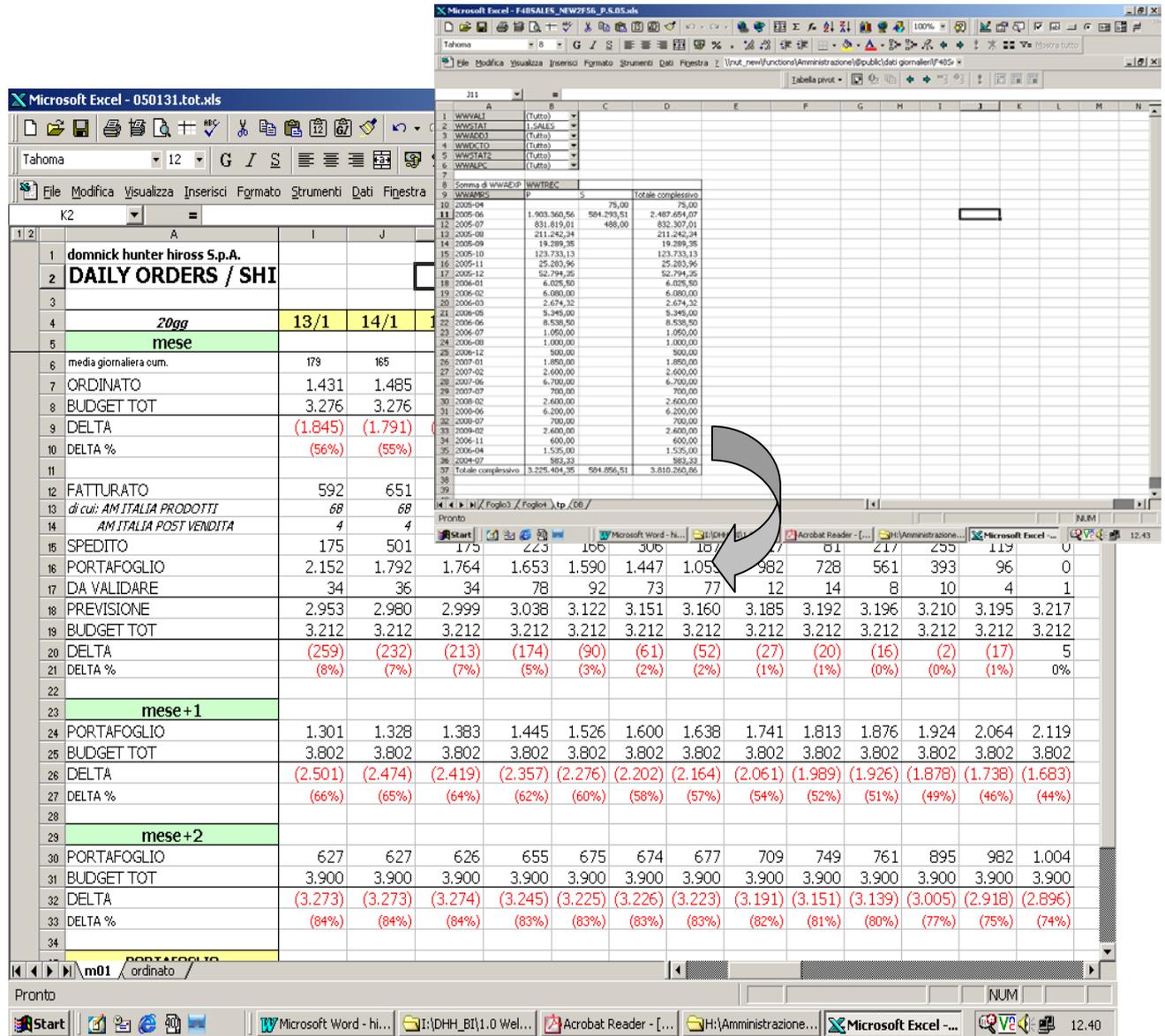


Figura 3: Daily report in Excel aggiornato manualmente

OutlookSoft Everest Application Set : SALES - SALES_act_bdg.XLT

File Modifica Visualizza Inserisci Formato Strumenti Dati Finestra eAnalyze eSubmit ePublish eTools ? \\nut_new\dh\Report\bi\sales\SALES_act_bdg.XLT

SALES_REPORTI ACCOUNT_MANAGER TOT_ ACCOUNTS FATT_VAL ADJUSTMENTS TOT_REV CATEGORY ACT_Fatt

J47

39 06/15/05 11:08 AM Server:HTTP://BI User:TAGLIABUE App:5 06/14/05 11:08 AM Server:HTTP://BI User:TAGLIABUE App:SALES Application:SALES_REPORTING

SALES

Total Revenues Current Period **Total Revenues Year to date**

Current Month				Current Month				Side	SELECT:	DESCRIPTION:
2005	2005	Batter/(Worra)		2004	2004	Batter/(Worra)				
Invoiced	Budget	Val	%	Invoiced	Budget	Val	%			
1596	3.605	(2.009)	-55,72%	4.200	(2.604)	-62,00%	1011	2005.TOTAL	Total Year 2005	
691	373	318	85,28%	390	302	77,38%	1012	100	1.AM	
200	402	(202)	-50,22%	369	(168)	-45,48%	1013	TOT_ST	Total Statistic Area	
(0)	(9)	9	96,64%	(43)	43	99,27%	1014	TOT_PROD	Total Product	
2.487	4.371	(1.883)	-43,09%	4.916	(2.428)	-49,40%	10	TOT_VE	Total Salesman	
2.202	5.945	(3.743)	-62,97%	4.871	(2.670)	-54,81%	3021	TOT_A_MGR	Total Account Manager	
2.202	5.945	(3.743)	-62,97%	4.871	(2.670)	-54,81%	30	TOT_REV	Total Revenue	
771	2.229	(1.458)	-65,46%	2.056	(1.285)	-62,49%	4031	FATT	Salar * 000	
78	396	(318)	-80,26%	279	(201)	-71,99%	4032			
60	163	(103)	-62,93%	124	(63)	-51,13%	4033			
8	83	(76)	-90,39%	55	(47)	-85,15%	4034			
34	-	34	90	(56)	-	-100,00%	4036			
5	-	5	4	1	14,00%	100,00%	PRODUCT.4037			
1	-	1	11	(10)	-92,69%	4038				
956	2.872	(1.915)	-66,69%	2.618	(1.662)	-63,47%	40			
69	-	69	-	69	-	69	5041			
69	-	69	-	69	-	69	PRODUCT.50			
534	1.383	(848)	-61,36%	1.381	(847)	-61,31%	2001			
474	1.138	(664)	-58,33%	1.075	(600)	-55,87%	2002			
168	803	(635)	-79,07%	438	(270)	-61,63%	2003			
49	47	2	3,52%	112	(63)	-56,37%	2004			
1.225	3.371	(2.145)	-63,65%	3.006	(1.788)	-59,23%	20			
6.940	16.558	(9.619)	-58,09%	15.411	(8.471)	-54,97%	TOT_PROD			

select dimension: **PRODOTT**

Year to Date				Year to Date			
2005	Year	Batter/(Worra)		2004	Batter/(Worra)		
Invoiced	Budget	Val	%	Invoiced	Val	%	
1596	3.605	(2.009)	-55,72%	4.200	(2.604)	-62,00%	
691	373	318	85,28%	390	302	77,38%	
200	402	(202)	-50,22%	369	(168)	-45,48%	
(0)	(9)	9	96,64%	(43)	43	99,27%	
2.487	4.371	(1.883)	-43,09%	4.916	(2.428)	-49,40%	
2.202	5.945	(3.743)	-62,97%	4.871	(2.670)	-54,81%	
2.202	5.945	(3.743)	-62,97%	4.871	(2.670)	-54,81%	
771	2.229	(1.458)	-65,46%	2.056	(1.285)	-62,49%	
78	396	(318)	-80,26%	279	(201)	-71,99%	
60	163	(103)	-62,93%	124	(63)	-51,13%	
8	83	(76)	-90,39%	55	(47)	-85,15%	
34	-	34	90	(56)	-	-100,00%	
5	-	5	4	1	14,00%		
1	-	1	11	(10)	-92,69%		
956	2.872	(1.915)	-66,69%	2.618	(1.662)	-63,47%	
69	-	69	-	69	-	69	
69	-	69	-	69	-	69	
534	1.383	(848)	-61,36%	1.381	(847)	-61,31%	
474	1.138	(664)	-58,33%	1.075	(600)	-55,87%	
168	803	(635)	-79,07%	438	(270)	-61,63%	
49	47	2	3,52%	112	(63)	-56,37%	
1.225	3.371	(2.145)	-63,65%	3.006	(1.788)	-59,23%	
6.940	16.558	(9.619)	-58,09%	15.411	(8.471)	-54,97%	

REPORT VIEW

Ready NUM

Microsoft Word OutlookSoft Everest ... 11.25

Figura 4: Report vendite navigabile in Everest Outlook Soft

CAPITOLO II

DOMNICK HUNTER HIROSS

Domnick hunter hiross SPA rappresenta quella parte di Hiross, fondata nel 1964 dal Sig. Rossi ed il Sig. Bianchi a S. Angelo di Piove di Sacco in provincia di Padova, dedicata allo sviluppo e alla produzione di prodotti per la filtrazione, purificazione e separazione dell'aria e dell'acqua: per mezzo di filtri, essiccatori e refrigeratori.

Nel 1999 viene acquistata dal gruppo britannico Domnick Hunter, presente nel mondo con filiali in più di quindici paesi, oggi Hiross è un'azienda con un fatturato di quasi 45 milioni €, in forte crescita (+ 20% 04/03), e con una buona percentuale di utili.

L'azienda opera in un mercato in forte espansione (refrigerazione) e con una gamma di prodotti che possono essere impiegati in molti settori (pompe di benzina, catene di montaggio, industria alimentare, ecc. ...).

Al fine di aumentare il proprio volume di affari, l'azienda persegue un obiettivo di globalizzazione, delle vendite e della produzione, oltre che un'ottimizzazione dei processi interni, iniziata nel 2003 con l'introduzione di un sistema di Controllo di Gestione e la misurazione delle performance con l'aiuto della Balanced Scorecard ed una serie di Key Performance Indicator (KPI).

Più in particolare, Domnick hunter hiross ha attraversato una fase di Business Process Re-engineering, in cui l'azienda – precedentemente organizzata per aree verticali – si è organizzata per processi

orizzontali ed ha attivato un processo di ristrutturazione dei processi nell'area del Controllo di Gestione.

Nel 2003 ha pubblicato il cosiddetto "navigator di gruppo" attraverso l'albero dei KPI aziendali. Per rendere automatizzato il calcolo di tali indicatori, attraverso un Business Performance Management, la scelta aziendale è ricaduta sullo strumento di business intelligence Outlooksoft Everest. La misurazione del rendimento aziendale attraverso la Balanced Scorecard è già in atto presso la casa madre inglese di Newcastle ed è stata, attualmente, adottata anche in Italia con l'impiego di una serie di indicatori (116 in totale) all'interno delle 4 prospettive di Norton e Kaplan:

- Finanziaria;
- Cliente;
- Processi interni;
- Innovazione ed apprendimento;

Domnick Hunter Hiross si distingue per l'orientamento all'innovazione dei prodotti, rilevanti investimenti in ricerca e sviluppo, flessibilità e il servizio al cliente, processi automatizzati in molte aree (es.: automazione dell'offerta tramite il portale B2B e vendite via portale B2B) e, soprattutto, attraverso un'organizzazione interna efficiente, monitorata con l'aiuto della Balanced Scorecard.

CAPITOLO III

PROGETTO "BUSINESS INTELLIGENCE"

Domnick Hunter Hiross ha attivato nel corso dell' anno 2003 un *business process reengineering* dell'area Controllo di Gestione con il duplice scopo di:

- Individuare i più importanti processi che lo interessavano;
- Migliorando l'efficacia e l'efficienza delle attività in carico alla funzione;

Per pervenire a tali obiettivi l'azienda ha deciso di avvalersi di uno strumento di data warehousing: Il Data Warehouse è una collezione orientata al soggetto, integrata, variante rispetto al tempo e non volatile, di dati in supporto al processo decisionale del management (*Bill Inmon*).

in modo da rendere automatiche molte delle attività volte alla raccolta, pubblicazione e analisi di informazioni la scelta è ricaduta sul prodotto Everest di Outlooksoft.

I progetti 2004-2005 coinvolgono le aree vendite, magazzino, acquisti, bilancio e produzione.

Il progetto nel complesso (figura 5) si articola su cinque Fatti o Cubi. Sono "FATTI" le variabili o misure, tipicamente di natura numerica, che sono al centro delle analisi del sistema di Business Intelligence. Nel caso in esame i Fatti sono: le Vendite, il Magazzino, il Budget, il Bilancio e la Produzione.

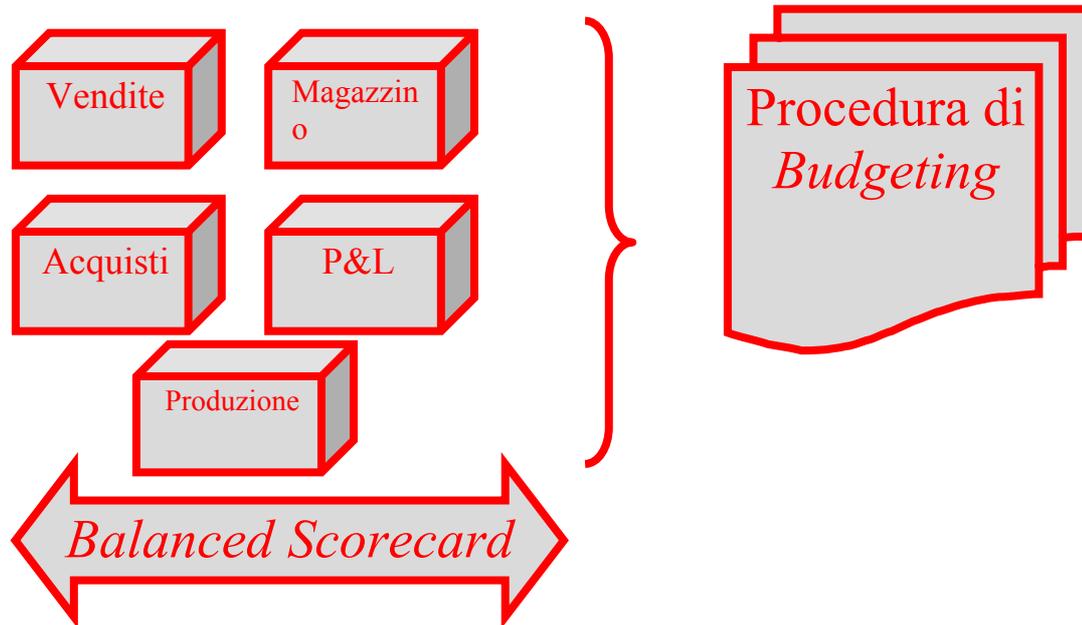


figura 5. Schema progetto BI

Ulteriore scopo del progetto Business Intelligence è di snellire le attività di reporting passando da report statici (figura 6) e aggiornati manualmente a report navigabili (figura 7) e aggiornati automaticamente in tempo reale che permettano operazioni di Roll-up e Drill-down fatte direttamente dai "consumatori di report".

I / area / product / range domnick hunter hiross S.p.A.
currency: euro/1000

			January 05		year to date				
Δ %	pr year	Δ %	all		actual	budget	Δ %	pr year	Δ %
27	556	5	STARLETTE		585	459	27	556	5
-	0	-	BLIZZARD		0	0	-	0	-
-	0	-	POLESTAR SMALL		52	0	-	0	-
3	666	14	POLESTAR MEDIUM		758	737	3	666	14

				January 05		year to date		
budget	Δ %	pr year	Δ %	AM: Italy		actual	budget	Δ %
25	14	27	4	STARLETTE		28	25	14
0	-	0	-	BLIZZARD		0	0	-
0	-	0	-					

sales per channel / area / product / range domnick hunter hiross S.p.A.

			January 05		year to date					
budget	Δ %	pr year	Δ %	IC: ROE		actual	budget	Δ %	pr year	Δ %
93	16	112	(4)	STARLETTE		107	93	16	156	3
0	-	0	-	BLIZZARD		0	0	-	0	-
0	-	0	-	POLESTAR SMALL		0	0	-	0	-
147	(1)	105	38	POLESTAR MEDIUM		145	147	(1)	161	111
134	38	104	78	POLESTAR BIG		185	134	38	0	0

figura 6. Report statici di Microsoft Excel

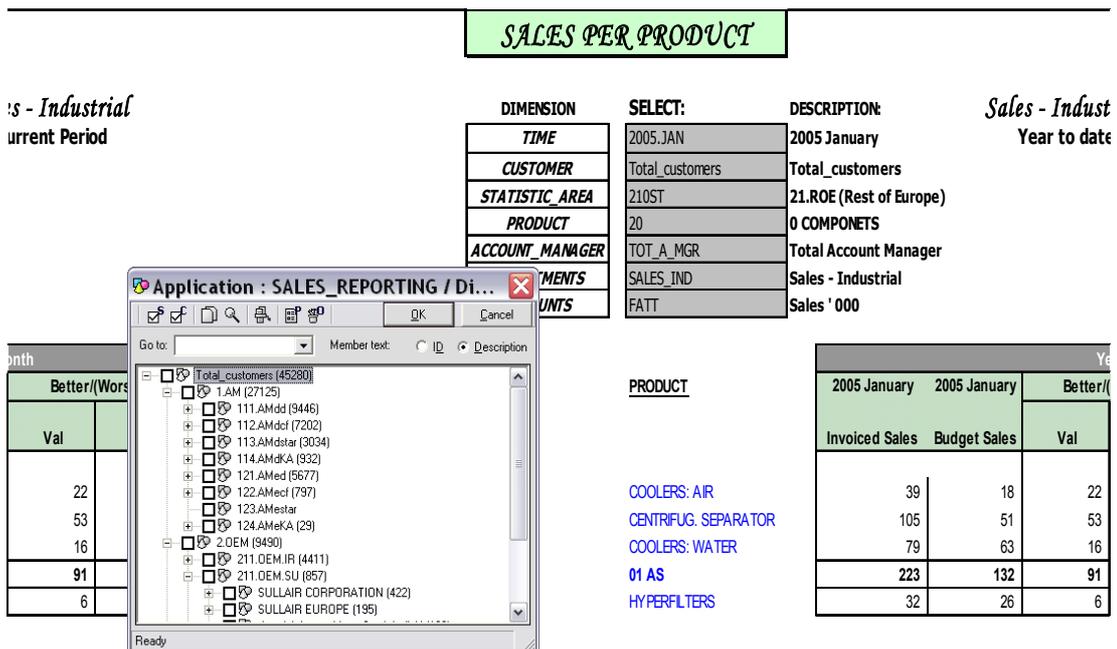


figura 7. report navigabile in Outlooksoft Everest

Le operazioni di Roll-up e Drill-down fanno parte dei tool OLAP che consentono operazioni di analisi e permettono all'utente di muoversi lungo la gerarchia di una dimensione passando a visione dei dati più o meno dettagliate. Nel caso della dimensione Tempo, per esempio, se si sta visualizzando i dati di un mese in particolare si potrà salire a un livello superiore e visualizzare un quarto di anno o i dati relativi a tutto l'anno. Quindi con le operazioni di Roll-up saliremo la gerarchia e con il Drill-down scenderemo lungo la stessa.

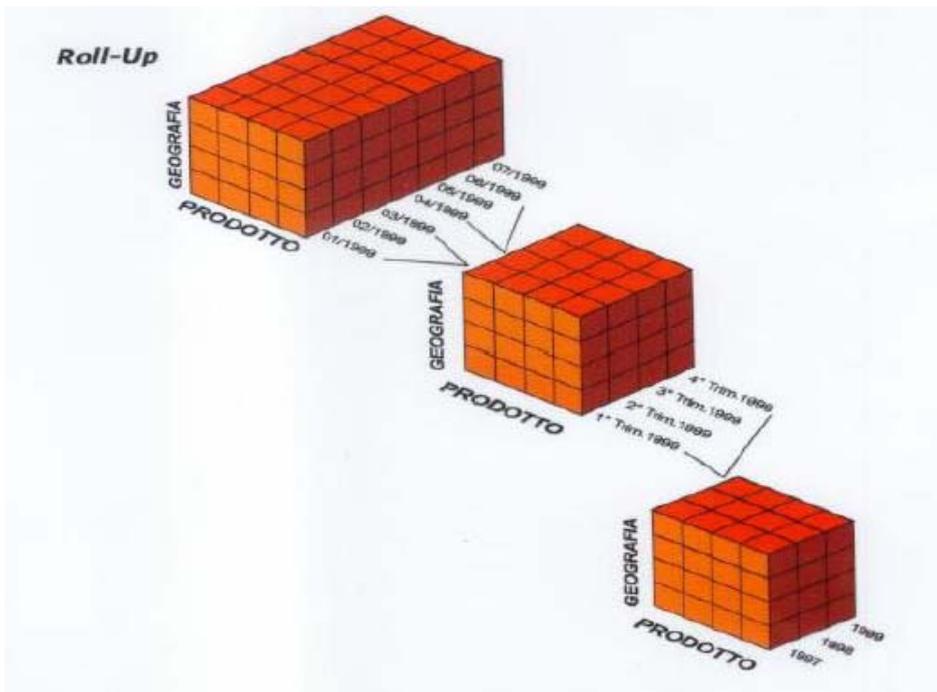


Figura 8. Roll-Up (Modelli e strutture per il DW, Dulli, Favero)

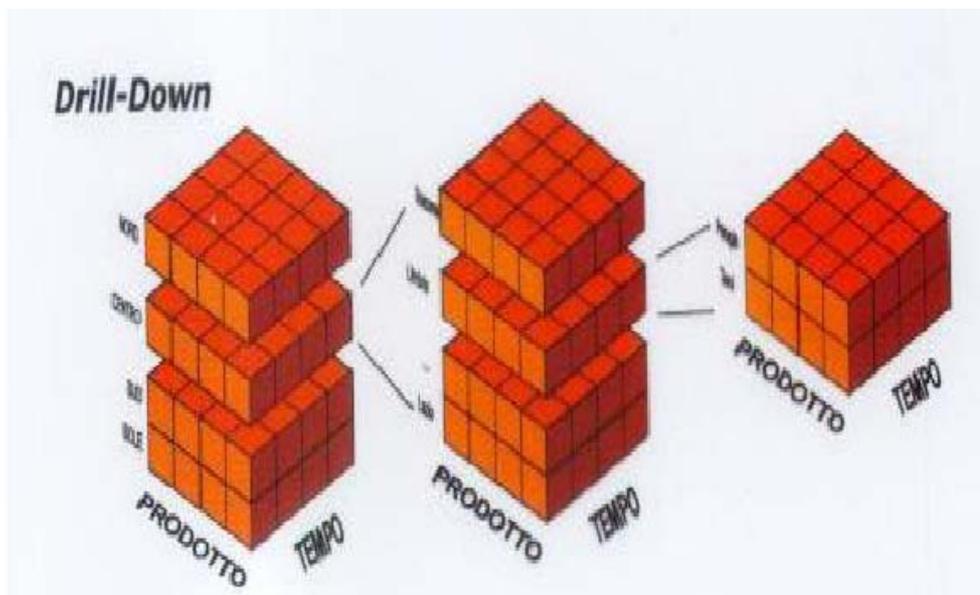


Figura 9. Drill-Down (Modelli e strutture per il DW, Dulli, Favero)

Ai livelli più alti di una gerarchia potremmo chiedere al tool OLAP un'altra operazione che permette la visione della somma (operazione di SUMMARIZATION), per periodo e prodotto, di tutte le vendite avvenute in Italia, per esempio.

L'operazione di ROTATION permette di ruotare gli assi del punto di vista delle dimensioni, in modo da poter ben mettere in evidenza le variabili oggetto di interesse.

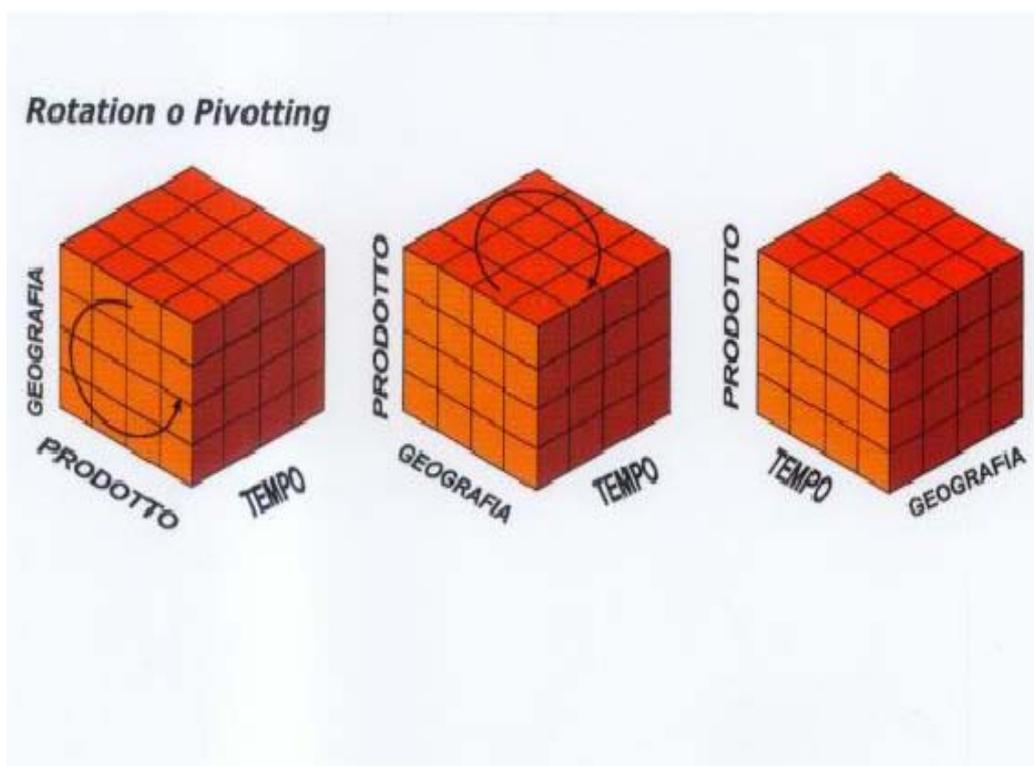


Figura 10. Rotation (Modelli e strutture per il DW, Dulli, Favero)

Slice&Dice (Figura 11), che significa affettare, permette di isolare una parte dei dati, e così effettuare analisi in particolari sottoinsiemi di informazioni.

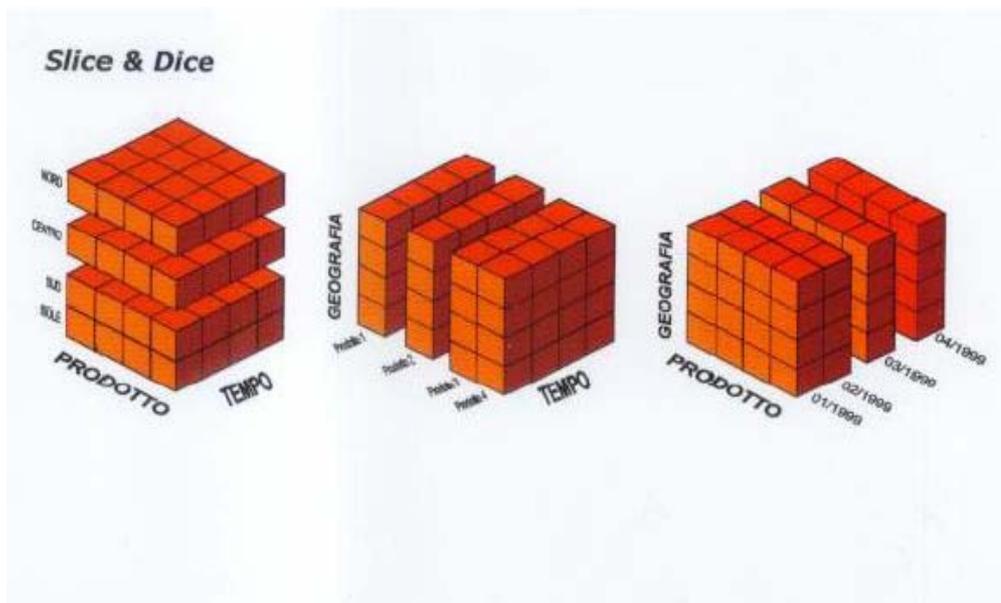


Figura 11. Slice&Dice (Modelli e strutture per il DW, Dulli, Favero)

A maggio 2005 le parti del progetto portate a termine sono il cubo VENDITE e il cubo MAGAZZINO, su quest'ultimo si stanno svolgendo training nei confronti degli utenti maggiormente coinvolti nella gestione e analisi del magazzino. Sono in fase di sviluppo le parti inerenti il bilancio (P&L) e gli ACQUISTI mentre il cubo che riguarderà la PRODUZIONE verrà implementato nell'ultima parte dell'anno 2005. Nei prossimi capitoli, dopo aver descritto il processo di Data Warehousing, si analizzerà nello specifico la struttura del Cubo VENDITE, primo dei progetti portato a termine.

CAPITOLO IV

IL PROCESSO DI DATA WAREHOUSING

Il Data Warehouse va considerato come l'obiettivo finale di un processo, formato da più step, tale processo è chiamato "Data Warehousing", formato da tre fasi portanti: Estrazione, Trasformazione (fisica e logica), e Trasporto dei dati.

4.1 DATI

L'attenzione in un Data Warehouse è posta sui soggetti principali del business aziendale. Attività importante è comprendere quali entità tenere sotto controllo e attraverso quali parametri. Tanto più i dati raccolti saranno *mirati al soggetto*, tanto più efficaci saranno le decisioni prese partendo da questa conoscenza.

L'*integrazione* dei dati è una delle attività che il Data Warehousing deve svolgere per riunire e riconciliare quelle informazioni che in molte organizzazioni sono disperse in sistemi informativi diversi e a volte indipendenti. Un livello minimo di integrazione è dato da:

- Convenzioni comuni per i nomi
- Uguali strutture di codifica
- Stessi attributi fisici dei dati
- Comuni criteri di misurazione delle variabili

Se una di queste condizioni non fosse verificata, i dati dovranno essere modificati e resi consistenti prima del caricamento nel Data Warehouse.

Questi dati sono, per loro natura, di tipo *storico*, infatti una delle maggiori esigenze che questa tecnologia permette di soddisfare, è la possibilità di analizzare le tendenze passate al fine di effettuare previsioni per il futuro.

Nel Data Warehouse i dati *non sono modificati* ma costantemente aggiunti alla base dati esistente, avremo quindi un insieme iniziale di dati caricato durante il primo popolamento del database e, su base periodica, un nuovo insieme di dati che andranno ad aggiungersi ai precedenti. Tali dati sono utilizzati al solo scopo di analisi.

All'interno delle aziende esistono di fatto due sistemi informativi diversi: gli "On Line Transaction Processing" che sovrintendono all'operatività giornaliera dell'azienda, e gli "On Line Analytical Processing" che soddisfano il bisogno di informazione attraverso estrazioni di dati, elaborazioni e report.



Figura 12. OLAP OLTP (Modelli e strutture per il DW, Dulli, Favero)

4.2 SISTEMI OLTP

I sistemi OLTP sono generalmente utilizzati per registrare, modificare e visualizzare singoli record in base alle esigenze operative delle varie funzioni aziendali. Questi sistemi pur custodendo al loro interno dati legati alla dimensione temporale, spesso non considerano dati storici di gestioni passate, di conseguenza il modo per ottenere tali informazioni è una serie di operazioni di "restore" da supporti di backup.

4.3 SISTEMI OLAP

Nei sistemi OLAP i dati sono molto statici, a differenza degli OLTP, una volta caricati non vengono modificati, mentre gli inserimenti di nuovi dati avvengono, normalmente, attraverso processi di tipo "batch".

Diventa quindi produttivo "denormalizzare" lo schema dei dati OLTP raggruppando le informazioni in una o poche tabelle per ridurre il costo delle operazioni di Join nella ricerca dei dati e avere quindi query più veloci.

La *semplicità di analisi* e la *velocità dell'estrazione* dei dati sono le due maggiori priorità dei modelli concettuali per l'analisi di tipo OLAP. La semplicità di analisi si ottiene concentrando l'attenzione sui soggetti principali e tangibili del business aziendale, "*FATTI*", e specificandone gli aspetti sui quali si vorrà poter svolgere delle analisi, "*DIMENSIONI*".

4.4 MAPPATURA

La mappatura è la fase che precede l'estrazione dei dati. Vengono determinati innanzitutto gli attributi e le misure, OLTP o da dati esterni, che descriveranno il soggetto del Data Warehouse. Viene poi stilata una lista dove avremo, attributo per attributo, database, il campo in cui avverrà la memorizzazione ed eventuali trasformazioni. Questo tipo di informazioni sono definite Metadati.

4.5 ESTRAZIONE DATI

In questa fase i dati di interesse per l'oggetto del Data Warehouse sono estratti dai sistemi, per lo più OLTP, dove si trovano e memorizzati in un database di servizio. Nella maggior parte dei casi i sistemi da cui sono estratti i dati sono diversi tra loro, per questo vengono create delle routine, utili anche per assecondare regole aziendali per la memorizzazione e dare una prima correzione dei dati. Il processo di estrazione può avvenire per mezzo di programmi implementati con linguaggi procedurali come COBOL, C++, ecc. o in SQL, attraverso un Gateway, in outsourcing avvalendosi di una ditta esterna, utilizzando tool per l'estrazione dei dati.

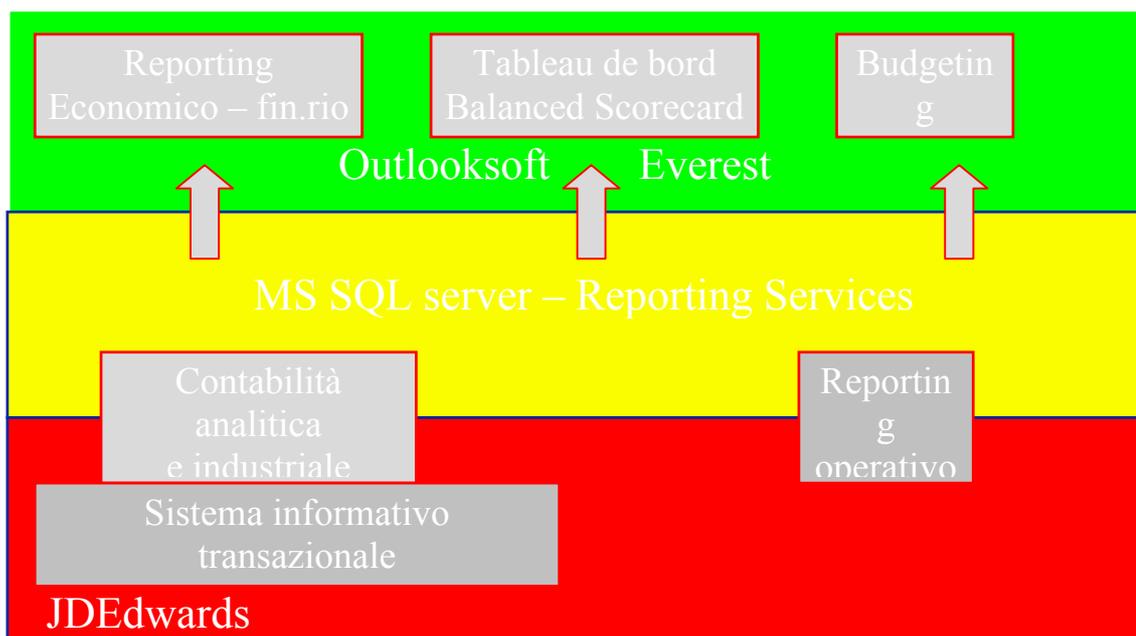


figura 13. architettura di riferimento progetto BI

4.6 AGGIORNAMENTO INCREMENTALE DEL DATA WAREHOUSE

L'aggiornamento dei dati all'interno del Data Warehouse è una fase delicata per ristrettezza di tempo, la difficoltà nel determinare quali dati modificati da caricare. Vi sono due tipi di metodi di aggiornamento: Statici e Dinamici.

4.6.1 METODI STATICI

Questi metodi si basano sulla registrazione di una snapshot dei dati OLTP in un particolare momento e il suo confronto con altre caricate in precedenza; in alcuni casi questo implica la registrazione di tutti i dati OLTP disponibili, mentre nella maggior parte dei casi solo una parte viene caricata ed elaborata.

4.6.2 METODI DINAMICI

I metodi dinamici sono applicati con successo quando i le modifiche avvenute nei dati OLTP da caricare sono relativamente poche rispetto la totalità dei dati stessi.

4.7 TRASFORMAZIONI FISICHE

Le trasformazioni fisiche sono volte alla omogeneità e alla "pulizia" dei dati. Le operazioni di trasformazione più comuni sono: Trasformazione dei termini OLTP in termini di uso comune, unificazione dei formati fisici di memorizzazione, controllo della consistenza nell'uso dei valori degli attributi, trattamento dei dati mancanti o errati.

4.8 TRASFORMAZIONI LOGICHE

Le trasformazioni logiche coinvolgono la struttura dei dati OLTP rendendoli idonei a rispondere al modello concettuale OLAP. La più conosciuta è la Denormalizzazione.

CAPITOLO V

PROGETTO "CUBO VENDITE"

5.1 I FATTI

Come già detto nel capitolo III sono "FATTI" le variabili o misure, tipicamente di natura numerica, che sono al centro delle analisi del sistema di Business Intelligence.

I database multi dimensionali organizzano le informazioni all'interno di strutture matriciali a più dimensioni; lo schema per questi database è facilmente rappresentabile per fatti aventi fino a tre dimensioni, ma il limite di dimensioni utilizzabili è dato solo dalla comprensibilità della struttura.

Modello Multidimensionale, cubi e ipercubi

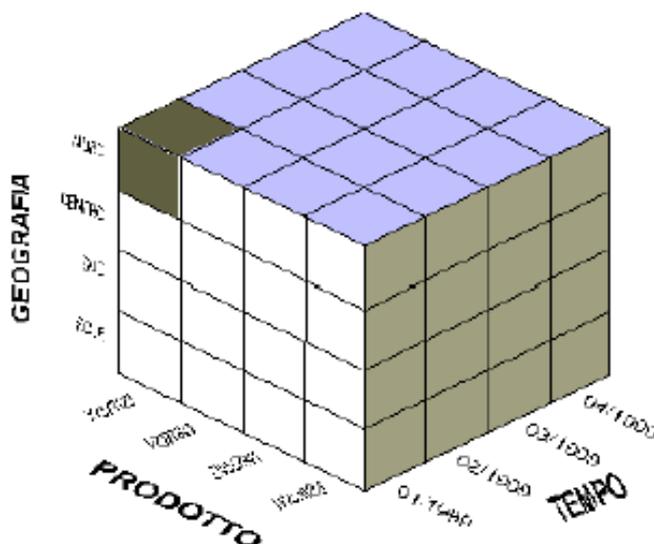


Figura 14. Multidimensional Model (Modelli e strutture per il DW, Dulli, Favero)

Nel Multi Dimensional Data Base, rappresentato in figura 14, ogni cella sarà caratterizzata dall'incrocio degli attributi lungo le tre dimensioni. Gli MDDB, mediante gli strumenti OLAP di cui dispongono, permettono di effettuare operazioni di analisi anche complesse in modo semplice e veloce grazie all'utilizzo di più matrici multi dimensionali ottenute preaggregando i dati in base alla visualizzazioni e analisi maggiormente usati. In questa sezione verrà osservata l'ideazione di quella parte del progetto Business Intelligence denominata Fatto o Cubo "VENDITE".

5.2 LE DIMENSIONI

Quando si considera un Fatto è importante verificare quali informazioni sono disponibili ed a quale dettaglio. Tali informazioni vengono raggruppate in "Dimensioni" che raccolgono quelli che saranno i criteri sui quali si potrà svolgere l'analisi dei dati nel Data Warehouse. Le dimensioni del Cubo VENDITE sono state scelte in base alle informazioni che l'azienda ha già a disposizione e i criteri sui quali si svolgeranno le analisi. Tali di dimensioni sono le seguenti:

- Account Manager
- Salesman
- Time
- Statitic Area
- Invoicing Area
- InputCurrency
- Customer
- Category
- Delivery Area
- Measures

- Tipo Dato

Il massimo dettaglio al quale i dati sono memorizzati è definito "granularità" della Dimensione.

il Dimensional Model

- Fatti (nelle celle)
- Dimensioni (gli spigoli)

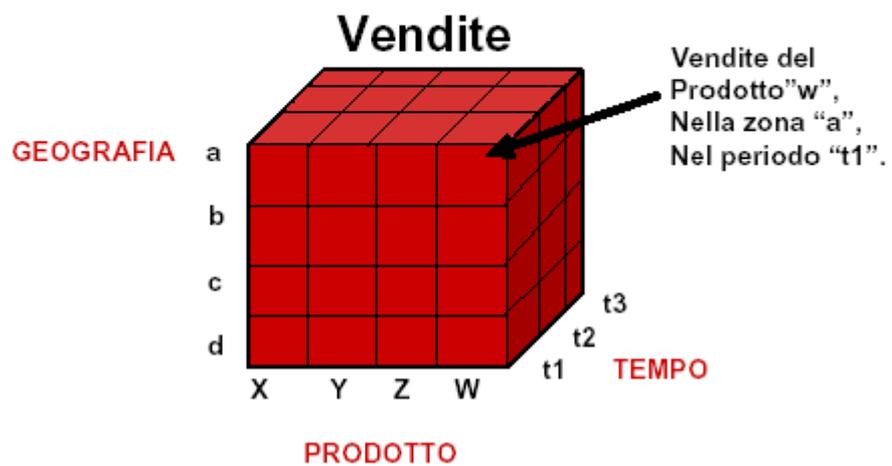


Figura 15. Dimensional Model (Modelli e strutture per il DW, Dulli, Favero)

5.3 GLI ATTRIBUTI

Le Dimensioni rappresentano concettualmente i Fatti e non sono implementate fisicamente nel Data Warehouse che avrà, invece, al suo interno gli attributi appartenenti ad ogni Dimensione, identificati dalle colonne del database. Prendendo in considerazione un esempio generico Vendite (figura 16) la dimensione geografica è rappresentata dagli attributi: "Area", "Regione", "Provincia". I valori assumibili da

ogni attributo saranno gli elementi sui quali si potranno effettuare analisi.

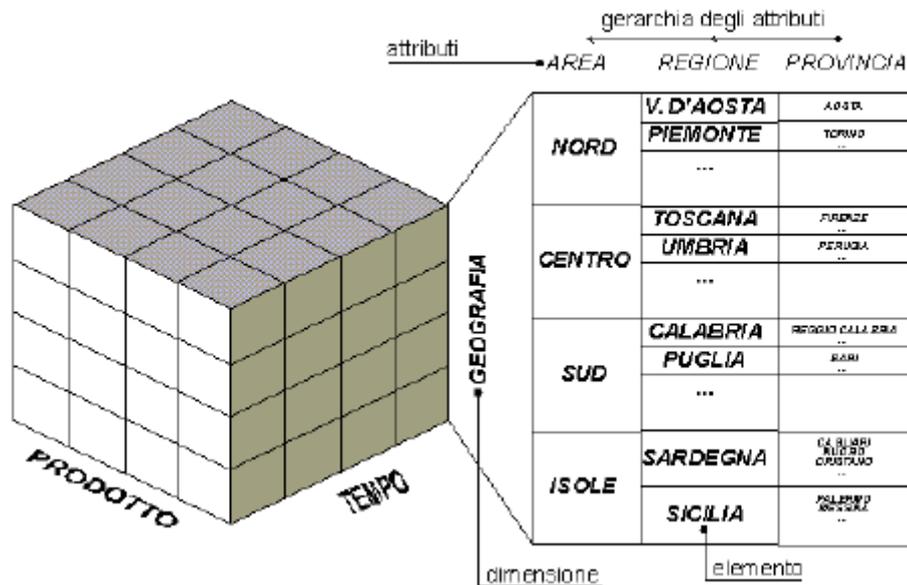


Figura 16. attributi (Modelli e strutture per il DW, Dulli)

5.4 GERARCHIE DEGLI ATTRIBUTI

Gli attributi di una dimensione sono classificati mediante *gerarchie* ben definite. Tali gerarchie possono avere più livelli. Esempio di gerarchia nella Dimensione Tempo può essere:

- Anno 2005;
 - Quarto 1;
 - Gennaio;
 - Febbraio;
 - 20050218;
 - Marzo;
 - Quarto 2;
 - Quarto 3;
 - Quarto 4;

5.5 ANALISI REQUISITI TECNICI

5.5.1 Naming degli oggetti del database

In questo paragrafo si danno le linee guida per il naming degli oggetti (tabelle, campi, ecc.) del database seguite durante la progettazione.

Al fine di semplificare lo sviluppo e la manutenzione del sistema sono da utilizzati nomi:

- semplici
- facili da ricordare
- facilmente interpretabili dall'utente

Nello sforzo di seguire questi postulati, è stato ritenuto necessario imporre delle regole sulla composizione dei nomi da assegnare agli oggetti durante la fase di progettazione.

Nella composizione dei nomi degli oggetti si sono utilizzati solo i seguenti caratteri alfanumerici:

lettere dell'alfabeto	A .. Z
Cifre	0 .. 9
Carattere underscore (utile per migliorare la leggibilità)	—

Tutti i caratteri alfanumerici non espressamente citati sono da ritenersi inammissibili.

5.5.2 Dettaglio Informativo

Per ogni dimensione è indicato:

- Contenuto
- Codifica elementi
- Struttura delle gerarchie
- Esempio della struttura gerarchica

5.6 DISEGNO LOGICO DELLA BASE DATI

5.6.1 Definizione delle dimensioni di analisi

- **Account Manager**

Contenuto

Indica il venditore associato ad un determinato cliente in anagrafica.

Codifica elementi

Gli elementi sono trattati con gli stessi codici utilizzati nell'anagrafica del sistema transazionale J.D.Edwards, con l'aggiunta di una "A" in coda agli stessi.

Struttura gerarchica

La dimensione contiene un'unica gerarchia a quattro livelli, totale incluso:

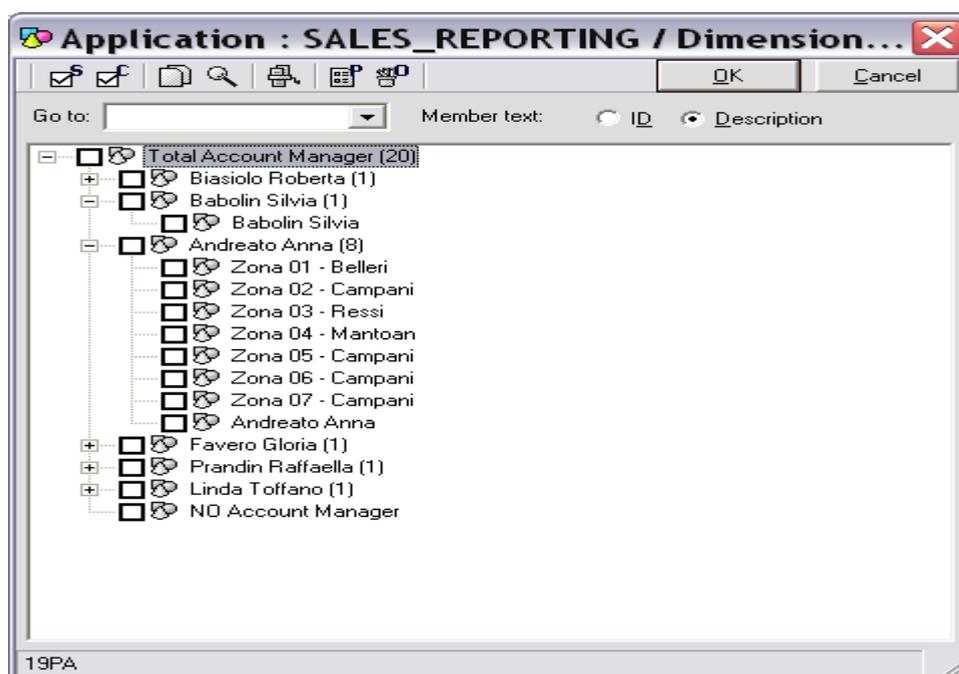
L0:TOTA - Total account manager

L1:09A-Biasiolo Roberta; 09B-Babolin Silvia; 09D-Andreato Anna;

L2:01-zona 01-Belleri; 02-zona 02-Campani; 03-zona 03-Ressi;.

L3:031-Cauteruccio; 045-Belleri; A12-Steri srl...

Schema della struttura gerarchica



- **Salesman**

Contenuto

Indica il venditore che ha effettuato la vendita il quale può differire dal venditore indicato in anagrafica per lo specifico cliente.

Codifica elementi

Gli elementi sono trattati con gli stessi codici utilizzati nell'anagrafica di J.D.Edwards, con l'aggiunta di una "E" in coda agli stessi.

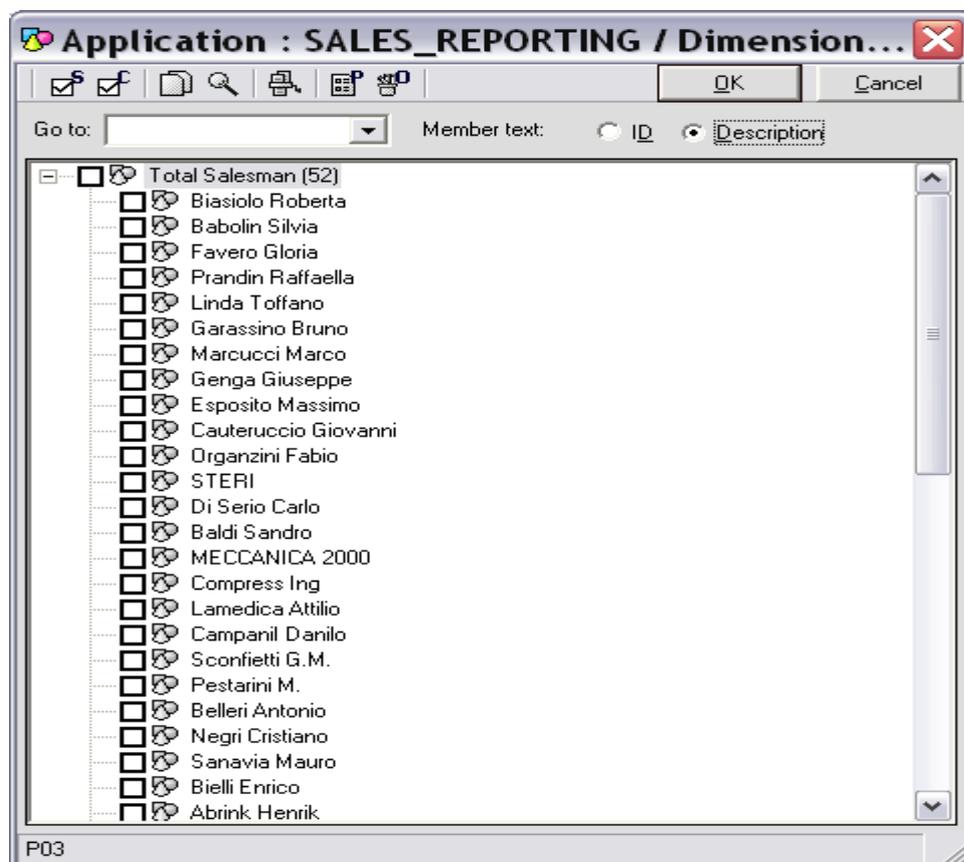
Struttura gerarchica

La dimensione contiene un'unica gerarchia a due livelli, totale incluso:

L0: TOTVE - Totale Salesman

L1: 09A-Biasiolo Roberta; 09B-Babolin Silvia; ;...

Schema della struttura gerarchica



- **Time**

Contenuto

Indica la dimensione tempo dell'applicazione, il livello di dettaglio con l'anno, la suddivisione in quarti dell'anno, il mese è il giorno.

Codifica elementi

Gli elementi sono rappresentati dai giorni dell'anno, codificati riportando anno, mese, giorno nel formato aaaammgg.

Struttura gerarchica

La dimensione contiene un'unica gerarchia a quattro livelli, totale incluso:

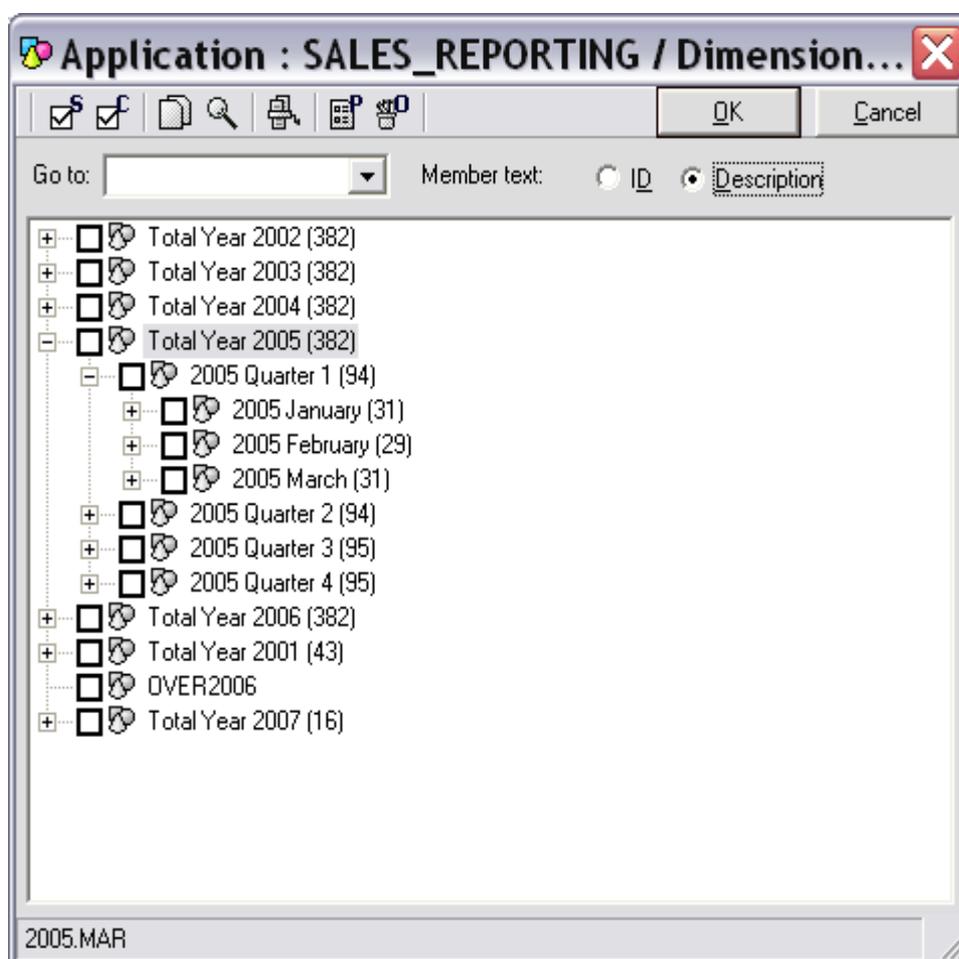
L0:2004.TOTAL–Total Year 2004; 2005.TOTAL–Total Year 2005;...

L1:2004.Q1–2004Quarter1; 2004.Q2–2004Quarter2;...

L2:2004.JAN – 2004 January; 2004.FEB – 2004 February;...

L3:20040101; 20040102; 20040103;...

Schema della struttura gerarchica



- **Product**

Contenuto

Riporta il codice prodotto oggetto della vendita.

Codifica elementi

La codifica è la stessa presente in J.D.Edwards. Avendo, alcuni codici prodotto, al loro interno il simbolo "\$" (carattere non consentito), questo è sostituito con la lettera "S".

Gli elementi gerarchici che rappresentano i totali, comuni a più gerarchie, mantengono lo stesso codice nelle diverse gerarchie. I codici totali, identici nel contenuto ma diversi per percorso di aggregazione, sono duplicati con l'aggiunta di un suffisso (es. 10 - Dryer in gerarchia Italia; 10 - UK in gerarchia UK)

Struttura delle gerarchie

La dimensione prodotto prevede cinque gerarchie di aggregazione:

Gerarchia Italia

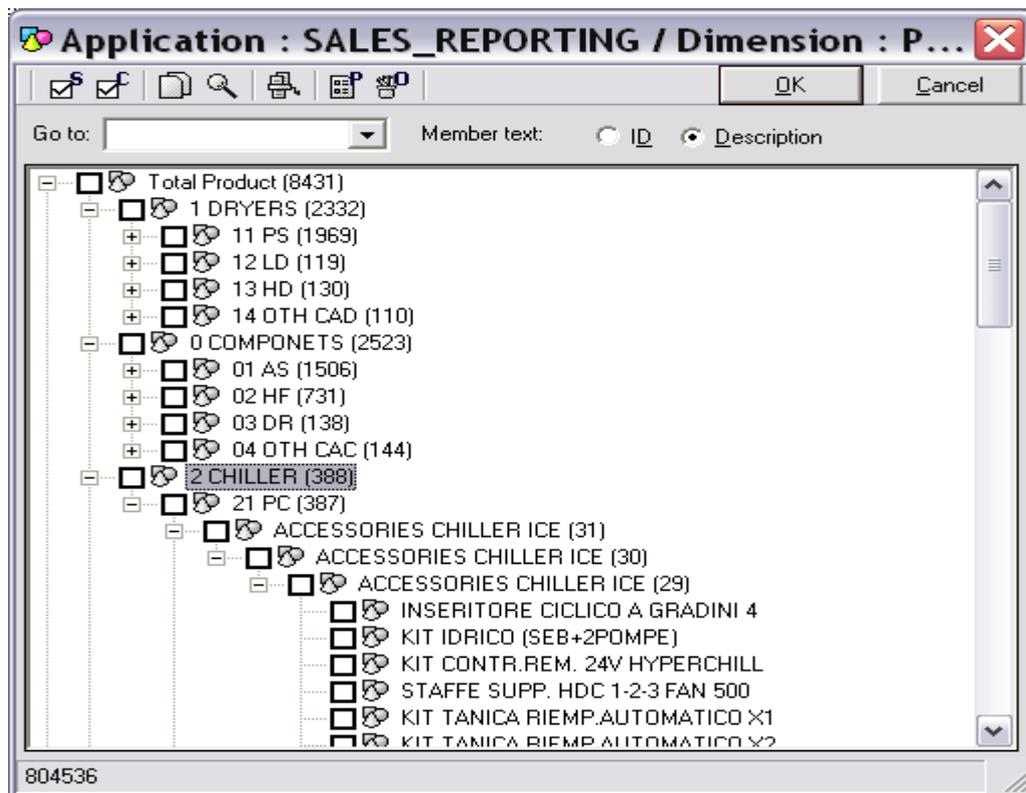
- Gerarchia UK
- Gerarchia per divisione
- Gerarchia per frequenza
- Gerarchia per brand

Struttura della gerarchia Italia

La gerarchia Italia è strutturata su sette livelli, totale incluso:

L0:TOTPROD-Totale prodotti	(Totale Italia)
L1:10 - 1.Dryer	(Linea di prodotto Italia)
L2:1001 - 01.PS	(Famiglia Italia)
L3:1A - POLESTAR ALTA PRESSIONE	(Range a 2 cifre)
L4:1AA - old POLAIR	(Range a 3 cifre)
L5:1AAA- old POLAIR	(Range a 4 cifre)
L6:805001 - POLAIR 004 H 50Bar 230	(Codice articolo)

Schema della struttura gerarchica Italia

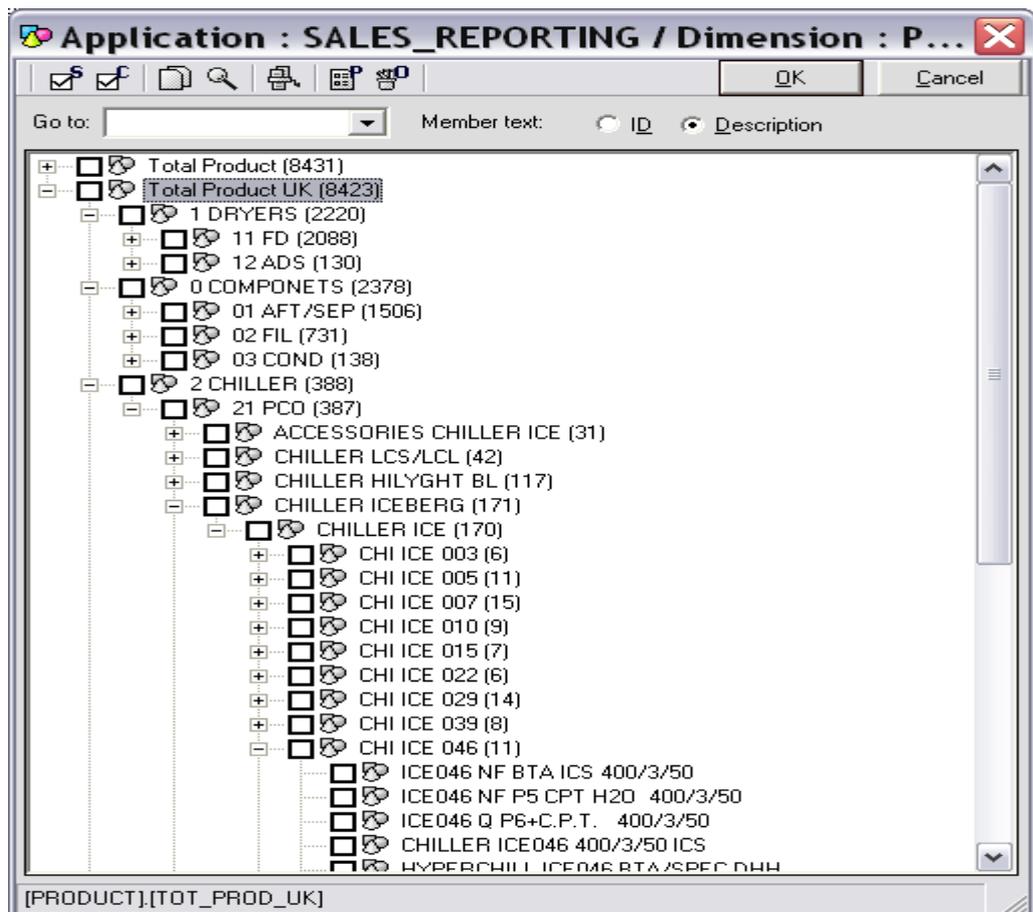


Struttura della gerarchia UK

La gerarchia UK è strutturata su otto livelli, totale incluso:

L0: TOTPRODUK – Totale prodotti UK	(Totale UK)
L1: 10UK - 1.Dryer	(Linea di prodotto UK)
L2: 1001UK – 01.FD	(Famiglia UK)
L2: 1001UK - 01.PS	(Famiglia Italia)
L3: 1A – POLESTAR ALTA PRESSIONE	(Range a 2 cifre)
L4: 1AA – old POLAIR	(Range a 3 cifre)
L5: 1AAA- old POLAIR	(Range a 4 cifre)
L6: 805001 – POLAIR 004 H 50Bar 230/1/50	(Codice articolo)

Schema della struttura gerarchica UK



Struttura della gerarchia per Divisione

La gerarchia per Divisione è strutturata su sei livelli, totale incluso:

- L0: TOTSTAB–Totale prodotti per stab. (Totale Stabilimento)
- L1: C0–C0.Stab. CAC: componenti (Divisione)
- L2: 0A – COOLERS WATER (Range a 2 cifre)
- L3: 0AA – COOLER FIX. TUBES (Range a 3 cifre)
- L4: 0AAA – COOLER FIX. TUBES (Range a 4 cifre)
- L5: 605051S20202 – WHF-035N/2B (Codice articolo)

Struttura della gerarchia per Frequenza

La gerarchia per Frequenza è strutturata su otto livelli, totale incluso:

L0:TOTPROD – Totale prodotti	(Totale Italia)
L1:10 - 1.Dryer	(Linea di prodotto Italia)
L2:1001 - 01.PS	(Famiglia Italia)
L3:1A-POLESTAR	(Range a 2 cifre)
L4:1A50Hz-POLESTAR 50Hz	(Range 2 cifre per freq.)
L4:1AA50Hz – old POLAIR.50Hz	(Range 3 cifre per freq.)
L5:1AAA50Hz- old POLAIR.50Hz	(Range 4 cifre per freq.)
L6: 805001 – POLAIR 004 H 50Bar	(Codice articolo)

Struttura della gerarchia per Brand

La gerarchia per Brand è strutturata su quattro livelli, totale incluso:

L0:TOT_BR – Totale prodotti per brand	(Totale prodotti)
L1:TOT_DH-Totale prodotti DH	(Totale Prod. Domnick Hunter)
L2:1NBA_DH – POLE*PGN 0040 DH	(Range 4 cifre per brand)
L3: 804211 – POLESTAR NEW040 DH	(Codice articolo)

- **Statistic Area**

Contenuto

Individua l'area geografica utilizzata per effettuare valutazioni di tipo statistico.

Codifica elementi

Sono stati utilizzati i codici presenti nell'anagrafica JDEdwards con l'aggiunta di un suffisso "ST" in coda al codice.

Struttura gerarchica

La dimensione contiene un'unica gerarchia a quattro livelli, totale incluso:

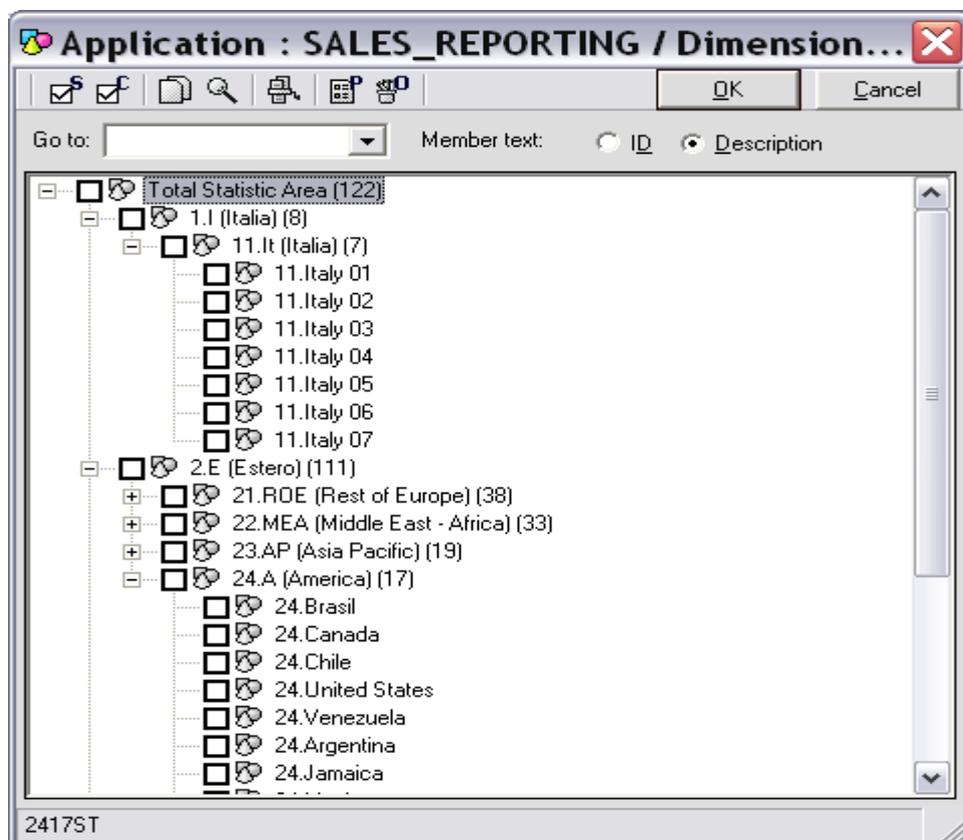
L0:TOTSTAT - Totale area statistica

L1:100ST - 1.I(Italia); 200ST - 2.E(Estero)

L2:210ST-21.ROE(Rest of Europe); 22.ST-22.MEA(Middle East Africa);

L3:2101ST-21.Germany; 2102ST-21.France; 2103ST-Spain;...

Schema della struttura gerarchica



- **Invoicing Area**

Contenuto

Individua l'area geografica nella quale il prodotto è effettivamente venduto.

Codifica elementi

Sono stati utilizzati i codici presenti nell'anagrafica JDEdwards.

Struttura gerarchica

La dimensione contiene un'unica gerarchia a cinque livelli, totale incluso:

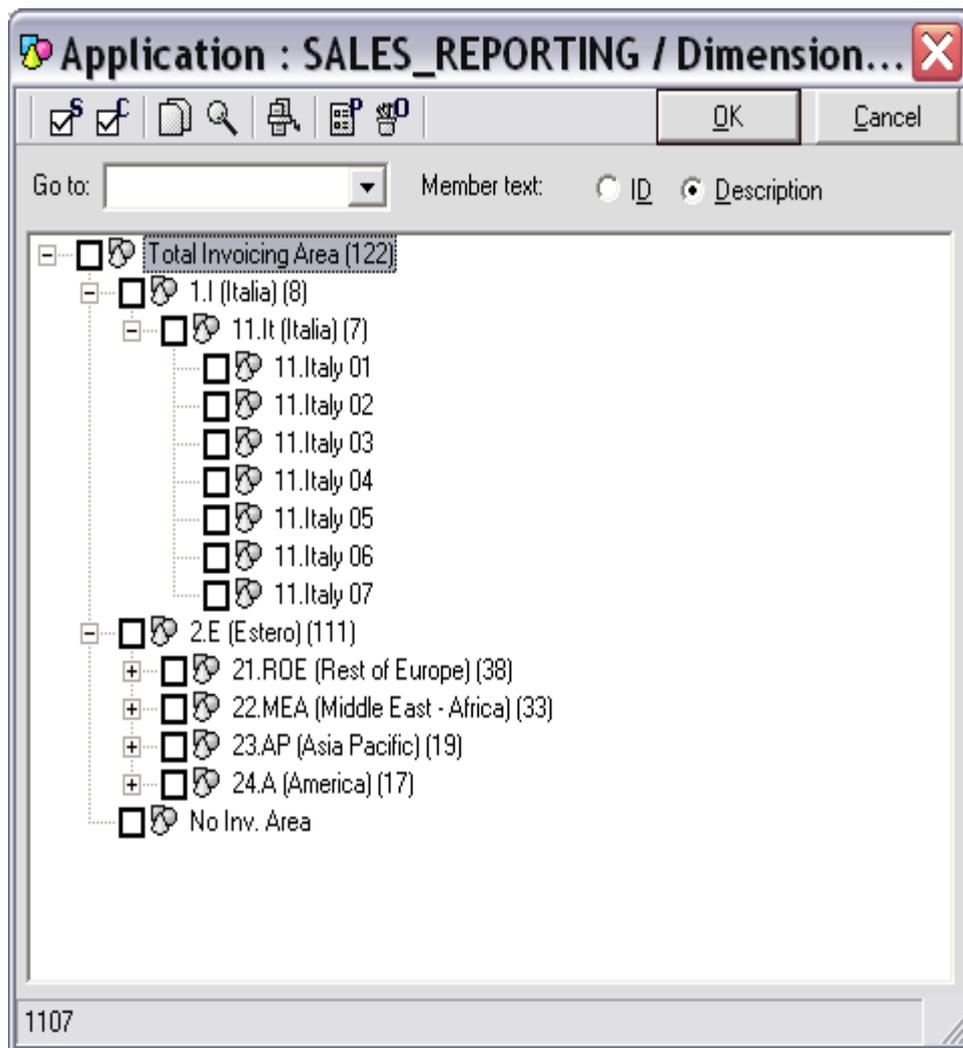
L0: TOT - Total invoicing area

L1: 100 - 1.I(Italia); 200 - 2.E(Estero)

L2: 210 - 21.ROE(Rest of Europe); 220 - 22.MEA(Middle East Africa);...

L3: 2101 - 21.Germany; 2102 - 21.France; 2103 - Spain;...

Schema della struttura gerarchica



- **Customer**

Contenuto

Rappresenta i clienti raggruppati per tre canali;

AM: After market

OEM: Sigla che indica grandi clienti

IC: Inter Company

Il dettaglio massimo è costituito dall'ordine di vendita riferito allo specifico cliente.

Codifica elementi

Sono stati utilizzati i codici presenti nell'anagrafica JDEdwards per il codice Canale, il codice Cliente e il numero d'Ordine .

Struttura gerarchica

La dimensione contiene un'unica gerarchia a quattro livelli, totale incluso:

L0: TOT – Total Customer

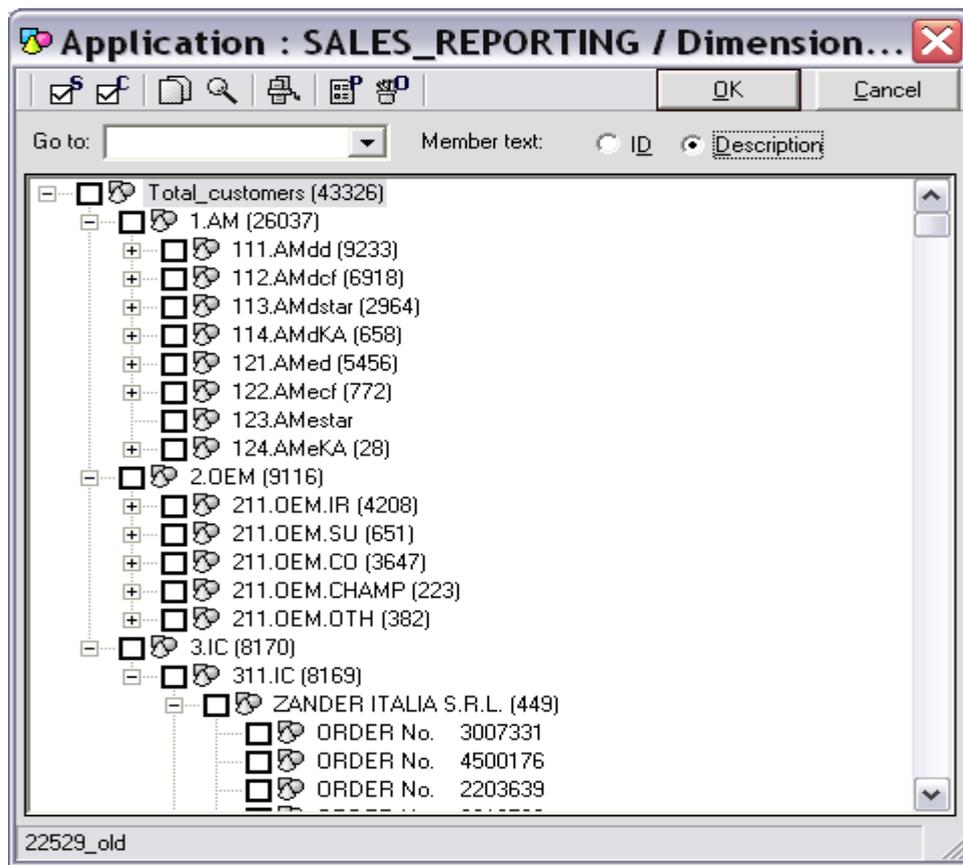
L1: 100 - 1.AM; 200 – 2.OEM; 300 – 3.IC

L2: 111 – 111 Amdd; 112 – 112 AMdcf ; 113 – Amed;...

L3: 132–Acqua minerale SanBenedetto; 140–ARIAC srl;...

L4: 3008496-Ordine 3008496; 4000232-Ordine 4000232;....

Schema della struttura gerarchica



- **Currency**

Contenuto

Rappresenta la valuta di riferimento utilizzata nel reporting: la fatturazione è riportata sia con la valuta di transazione sia in valuta di reporting ovvero Euro.

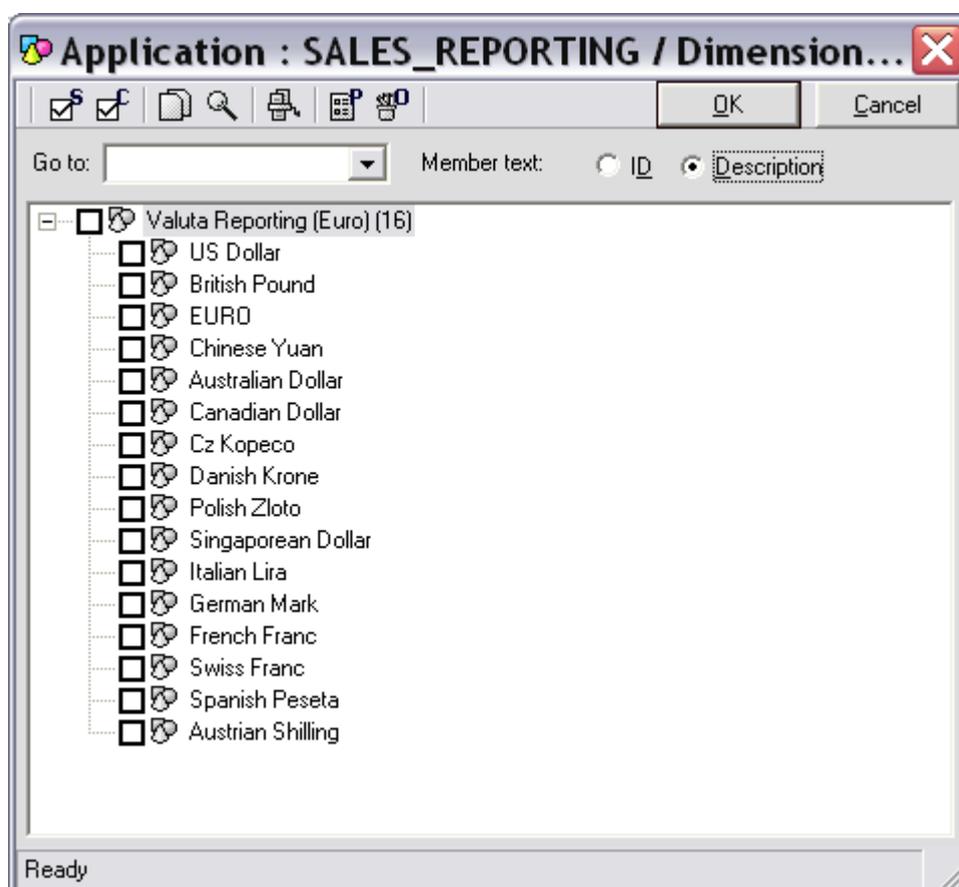
Codifica elementi

Sono stati utilizzati i codici presenti nell'anagrafica JDEdwards.

Struttura gerarchica

La dimensione non contiene gerarchie.

Schema della struttura gerarchica



- **Measures**

Contenuto

Indica i codici di conto inseriti nel Piano dei Conti.

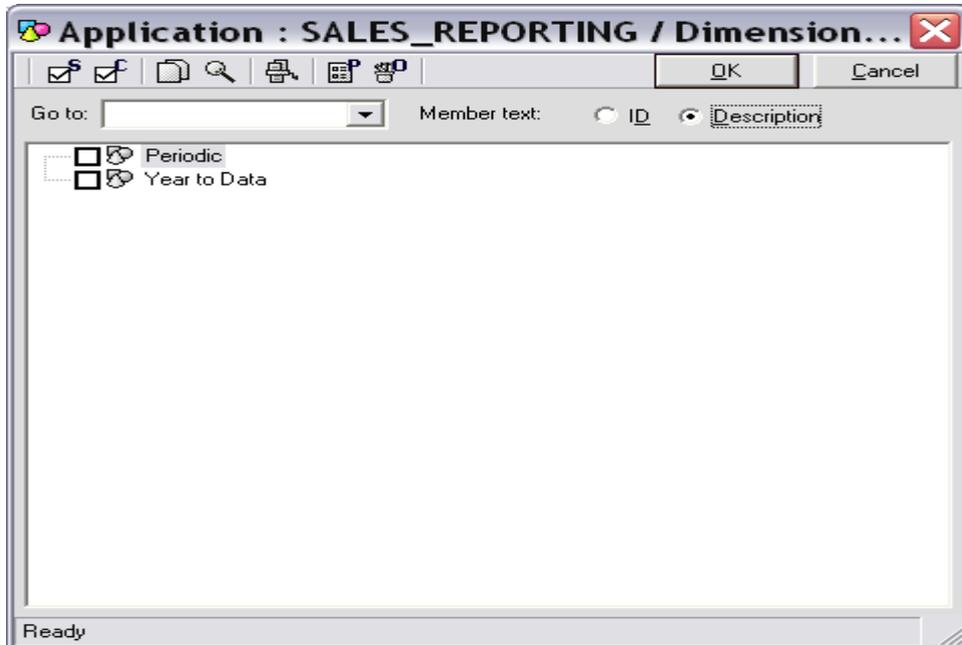
Codifica elementi

I codici riprendo in parte la descrizione degli elementi.

Struttura gerarchica

La dimensione non contiene gerarchie:

Schema della struttura gerarchica



- **Delivery Area**

Contenuto

Individua l'area geografica nella quale la merce viene spedita.

Codifica elementi

Sono stati utilizzati i codici presenti nell'anagrafica JDEdwards con l'aggiunta del un suffisso "SP" in coda al codice.

Struttura gerarchica

La dimensione contiene un'unica gerarchia a cinque livelli, totale incluso:

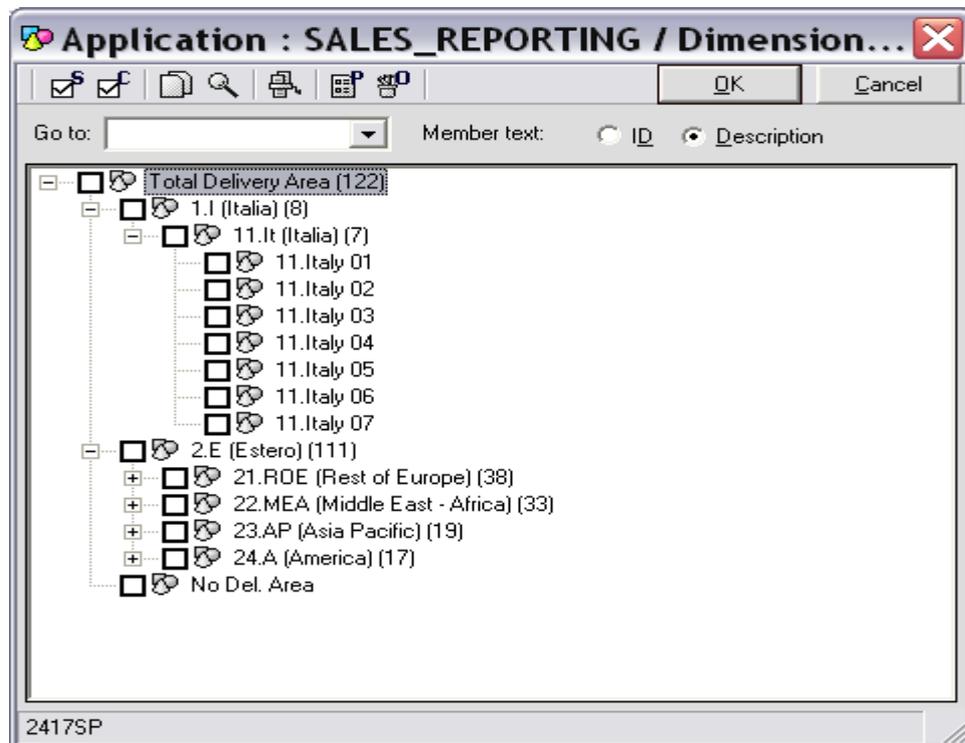
L0: TOTSPED - Totale paese di spedizione

L1: 100 - 1.I(Italia); 200 - 2.E(Estero)

L2: 210 - 21.ROE(Rest of Europe); 220 - 22.MEA(Middle East Africa);.....

L3: 2101 - 21.Germany; 2102 - 21.France; 2103 - Spain;.....

Schema della struttura gerarchica



- **Tipo dato**

Contenuto

Distingue il dato come importato dal sistema transazionale da eventuali rettifiche correzioni apportate per giungere al totale validato.

Codifica elementi

Sono stati utilizzati i codici definiti dai programmatori.

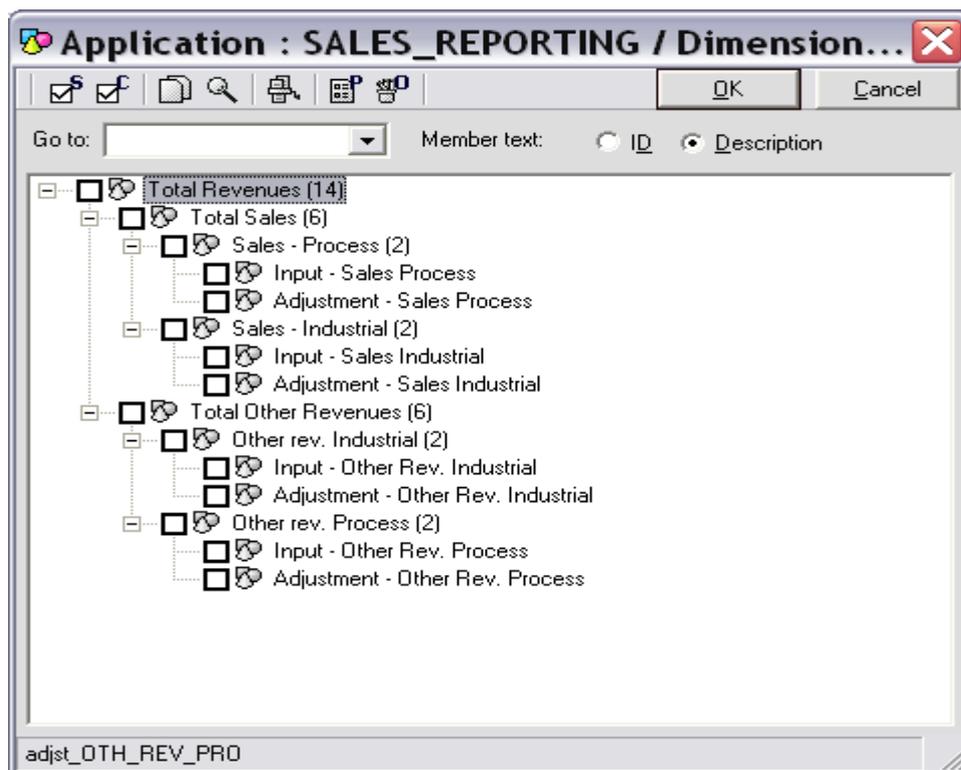
Struttura gerarchica

La dimensione contiene un'unica gerarchia a due livelli, totale incluso:

L0: Totale_validato – Totale_validato

L1: input – Dato importato; Adjustment – Rettifica

Schema della struttura gerarchica



5.6.2 LAYOUT MULTIDIMENSIONALE

Il seguente schema riassume tutte le possibili combinazioni delle dimensioni analizzate.

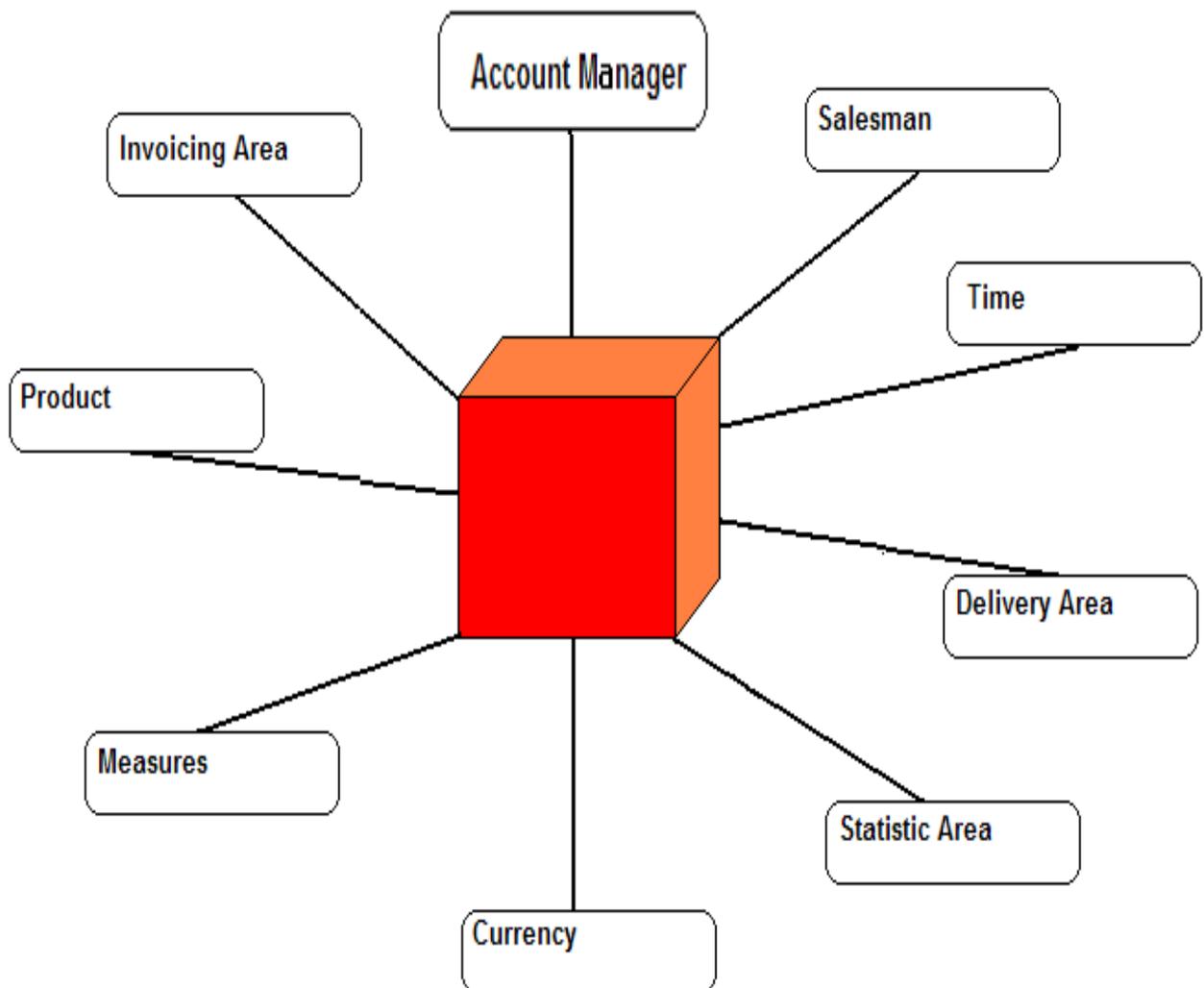


Figura 17. Layout multi dimensionale

CONCLUSIONI

L'introduzione di strumenti avanzati e di tecniche adeguate è sicuramente un processo arduo da realizzare, ma che fornisce un supporto importante all'attività di ogni realtà aziendale.

Il progetto Business Intelligence, se attuato con le giuste metodologie, consente all'alta direzione di prendere decisioni con una consapevolezza maggiore, consentendo un migliore utilizzo di tutte le risorse umane, finanziarie e di know-how a disposizione. La conoscenza dei dati consente la comprensione dei problemi e fornisce la base di supporto necessaria al dialogo e al coordinamento tra il management e gli altri "professionisti" dell'azienda.

L'obiettivo del management aziendale di DH Hiross, è quello di fornire nuovi strumenti per guidare meglio le scelte dei vari responsabili, e mettere a loro disposizione le migliori informazioni.

Per quanto riguarda l'aspetto statistico, il controllo di gestione necessita di processi di raccolta, analisi e interpretazione dei dati che scaturiscono dall'attività aziendale, il Data Warehouse si dà l'obiettivo di rendere queste operazioni più efficienti.

Durante lo sviluppo del progetto in questione le maggiori problematiche si sono manifestate nella difficoltà di intersecare le esigenze aziendali con i limiti di implementazione del nuovo software: problemi si sono riscontrati nella scelta della granularità di alcune dimensioni, come la Customer dove in alcuni casi avrebbe fatto comodo un diverso tipo di aggregazione dei dati: la suddivisione dei clienti nelle macro aree AM, OEM e IC, per la creazione di alcuni report che richiedono diversi criteri di associazione, risulta essere obsoleta se non d'intralcio, altre criticità si manifestano nella fase di aggiornamento del sistema, che per il cubo Vendite avviene

giornalmente: le cause sono da ricondursi alla difficoltà di caricamento dei dati non definiti in ogni loro parametro.

Nel momento in cui il nuovo sistema è stato reso disponibile alle varie aree aziendali, i componenti delle stesse hanno dovuto prendere confidenza con la nuova struttura, più complessa e dinamica, entrare nell'ottica della multidimensionalità e navigabilità del prodotto di Outlook soft, e con difficoltà "slegarsi" dalle vecchie procedure, meno complesse ma più rigide.

BIBLIOGRAFIA

- Bracchi "Processi aziendali e sistemi informativi", Franco Angeli;
- Dulli, Favero "Modelli e strutture per il Data Warehousing", Diade 2000;
- Bill Inmon "Building the Data Warehouse", John Wiley & Son, New York 1996;
- Ralph Kimball "The Data Warehouse LifeCycle Toolkit", Wiley 1996;
- Mentens ed altri "Informatica Aziendale", McGraw Hill;
- Quaderni di Management, n° 6 novembre-dicembre 2003;