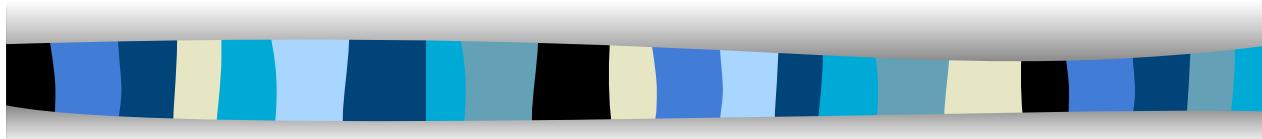


Il Data Warehousing



Dai dati alle informazioni

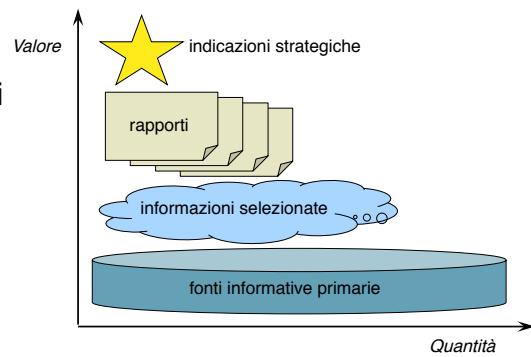
- L'informazione è un bene a valore crescente, necessario per pianificare e controllare le attività aziendali con efficacia
- Essa costituisce la materia prima che viene trasformata dai sistemi informativi, come i semilavorati vengono trasformati dai sistemi di produzione

~~dati = informazione~~

- Spesso la disponibilità di troppi dati rende arduo, se non impossibile, estrarre le informazioni veramente importanti

Dai dati alle informazioni

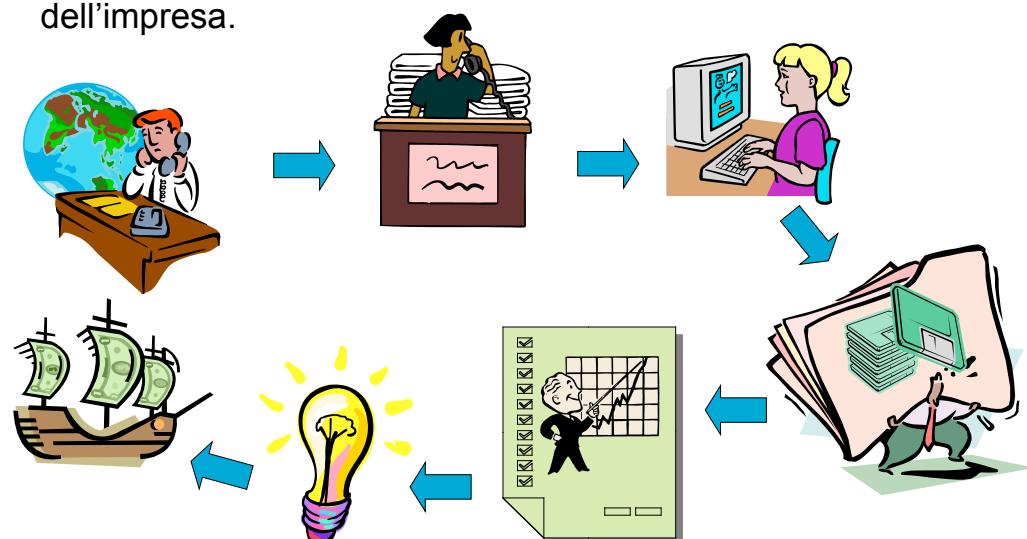
- Per ogni azienda è fondamentale poter disporre in maniera rapida e completa delle informazioni necessarie al processo decisionale: le indicazioni strategiche sono estrapolate principalmente dalla mole dei dati operazionali contenuti nei database aziendali, attraverso un procedimento di selezione e sintesi progressiva



3

Uno scenario tipico...

- .. è quello di una grande azienda, con numerose filiali, i cui dirigenti desiderano quantificare e valutare il contributo dato da ciascuna di esse al rendimento commerciale globale dell'impresa.



4

Uno scenario tipico...

- .. è quello di una grande azienda, con numerose filiali, i cui dirigenti desiderano quantificare e valutare il contributo dato da ciascuna di esse al rendimento commerciale globale dell'impresa.



5

Le interrogazioni

- **OLTP (On-Line Transactional Processing):**
 - ✓ Le interrogazioni eseguono transazioni che leggono e scrivono un ridotto numero di record da diverse tabelle legate da semplici relazioni
 - ✓ Il nucleo sostanziale del carico di lavoro è “congelato” all’interno dei programmi applicativi
- **OLAP (On-Line Analytical Processing):**
 - ✓ Le interrogazioni effettuano un’analisi multidimensionale che richiede la scansione di un’enorme quantità di record per calcolare un insieme di dati numerici di sintesi che quantifichino le prestazioni dell’azienda
 - ✓ L’interattività è una caratteristica irrinunciabile delle sessioni di analisi e fa sì che il carico di lavoro effettivo vari continuamente nel tempo

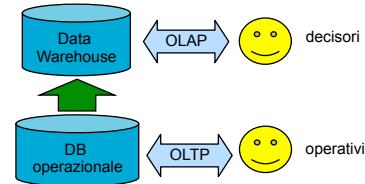
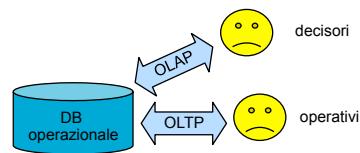
6

OLTP e OLAP

- Mescolare interrogazioni “analitiche” e “transazionali” di routine porta a inevitabili rallentamenti che rendono insoddisfatti gli utenti di entrambe le categorie



separare l’elaborazione di tipo analitico (OLAP) da quella legata alle transazioni (OLTP), costruendo un nuovo raccoglitore di informazioni che integri i dati provenienti da sorgenti di varia natura, li organizzi in una forma appropriata e li renda disponibili per scopi di analisi e valutazione finalizzate alla pianificazione e al processo decisionale



7

Alcune aree di utilità

- Commercio** (analisi delle vendite e dei reclami, controllo di spedizioni e inventari, cura del rapporto con i clienti)
- Manifattura** (controllo dei costi di produzione, supporto fornitori e ordini)
- Servizi finanziari** (analisi del rischio e delle carte di credito, rivelazione di frodi)
- Trasporti** (gestione parco mezzi)
- Telecomunicazioni** (analisi del flusso delle chiamate e del profilo dei clienti)
- Sanità** (analisi di ricoveri e dimissioni, contabilità per centri di costo)
-

8

Data Warehousing:

- Una collezione di metodi, tecnologie e strumenti di ausilio al *knowledge worker* (dirigente, amministratore, gestore, analista) per condurre analisi dei dati finalizzate all'attuazione di processi decisionali e al miglioramento del patrimonio informativo.

9

Le lamentele

- ☞ *abbiamo montagne di dati ma non possiamo accedervi!*
- ☞ *come è possibile che persone che svolgono lo stesso ruolo presentino risultati sostanzialmente diversi?*
- ☞ *vogliamo selezionare, raggruppare e manipolare i dati in ogni modo possibile!*
- ☞ *mostratemi solo ciò che è importante!*
- ☞ *tutti sanno che alcuni dati non sono corretti!*

R. Kimball, The Data Warehouse Toolkit



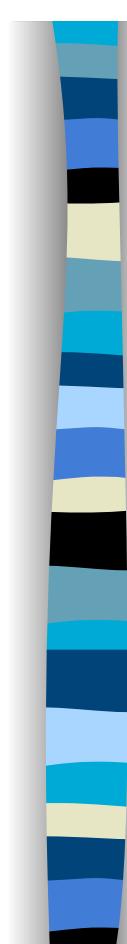
10



Caratteristiche del processo di warehousing

- **accessibilità** a utenti con conoscenze limitate di informatica e strutture dati;
- **integrazione dei dati** sulla base di un modello standard dell'impresa;
- **flessibilità di interrogazione** per trarre il massimo vantaggio dal patrimonio informativo esistente;
- **sintesi** per permettere analisi mirate ed efficaci;
- **rappresentazione multidimensionale** per offrire all'utente una visione intuitiva ed efficacemente manipolabile delle informazioni;
- **correttezza e completezza** dei dati integrati.

11



Il Data Warehouse

- Al centro del processo, il data warehouse è un contenitore di dati che si fa garante dei requisiti esposti.

➤ *Un Data Warehouse è una collezione di dati di supporto per il processo decisionale che presenta le seguenti caratteristiche:*

- ✓ *è orientata ai soggetti di interesse;*
- ✓ *è integrata e consistente;*
- ✓ *è rappresentativa dell'evoluzione temporale;*
- ✓ *non volatile.*

12

...orientato ai soggetti



13

...integrato e consistente

Il DW si appoggia a più fonti di dati eterogenee: dati estratti dall'ambiente di produzione, e quindi originariamente archiviati in basi di dati aziendali, o addirittura provenienti da sistemi informativi esterni all'azienda. Di tutti questi dati il DW restituisce una visione unificata.



14

...rappresentativo dell'evoluzione temporale

DB operazionali



Contenuto storico limitato,
spesso il tempo non è parte
delle chiavi,
i dati sono soggetti
ad aggiornamenti

DW

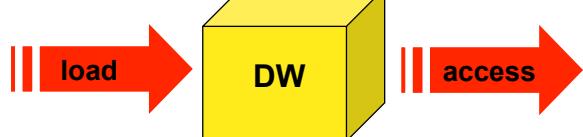
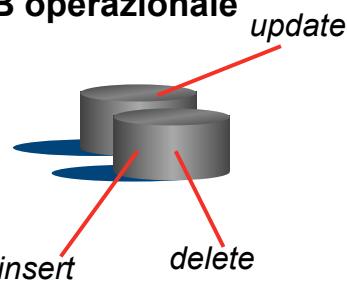


Ricco contenuto storico,
il tempo è parte delle chiavi,
una fotografia del dato
a un certo istante di tempo
non può essere aggiornata

15

...non volatile

DB operazionale



Grandi volumi di dati:
da qualche GB a diversi TB
in pochi anni

- ✓ in un DW non sono necessarie tecniche sofisticate di gestione delle transazioni come invece richiesto da un sistema operazionale
- ✓ problema chiave diventa il query-throughput

16

Riassumendo:

	Database operazionali	Data warehouse
utenti	migliaia	centinaia
carico di lavoro	transazioni predefinite	interrogazioni di analisi <i>ad hoc</i>
accesso	a centinaia di record, in lettura e scrittura	a milioni di record, per lo più in lettura
scopo	dipende dall'applicazione	supporto alle decisioni
dati	elementari, sia numerici sia alfanumerici	di sintesi, prevalentemente numerici
integrazione dei dati	per applicazione	per soggetto
qualità	in termini di integrità	in termini di consistenza
copertura temporale	solo dati correnti	dati correnti e storici
aggiornamenti	continui	periodici
modello	normalizzato	multidimensionale
ottimizzazione	per accessi OLTP su una frazione del database	per accessi OLAP su gran parte del database

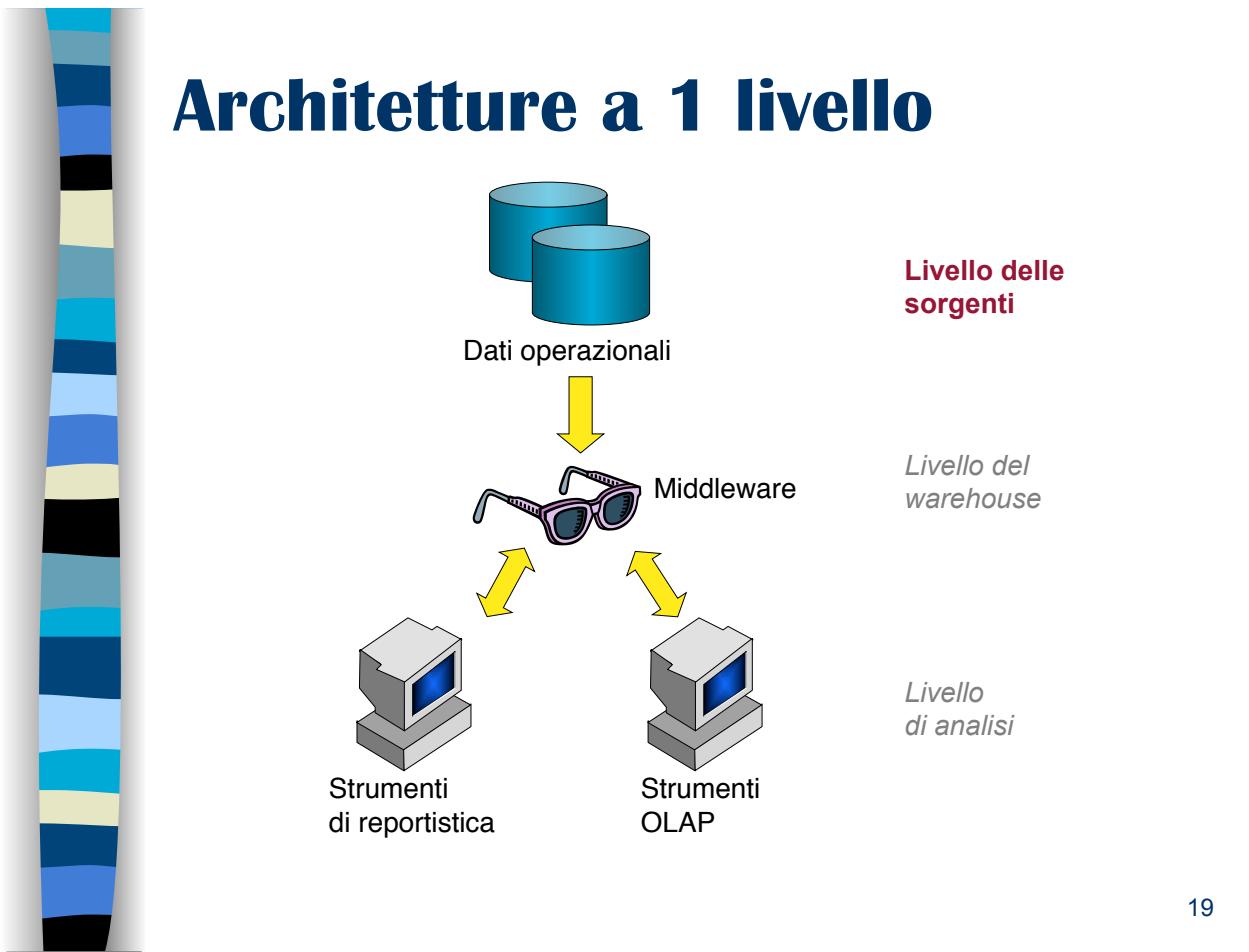
17

Architetture: requisiti

- ✓ **Separazione:** l'elaborazione analitica e quella transazionale devono essere mantenute il più possibile separate.
- ✓ **Scalabilità:** l'architettura hardware e software deve poter essere facilmente ridimensionata a fronte della crescita nel tempo dei volumi di dati da gestire ed elaborare e del numero di utenti da soddisfare.
- ✓ **Estendibilità:** deve essere possibile accogliere nuove applicazioni e tecnologie senza riprogettare integralmente il sistema.
- ✓ **Sicurezza:** il controllo sugli accessi è essenziale a causa della natura strategica dei dati memorizzati.
- ✓ **Amministrabilità:** la complessità dell'attività di amministrazione non deve risultare eccessiva.

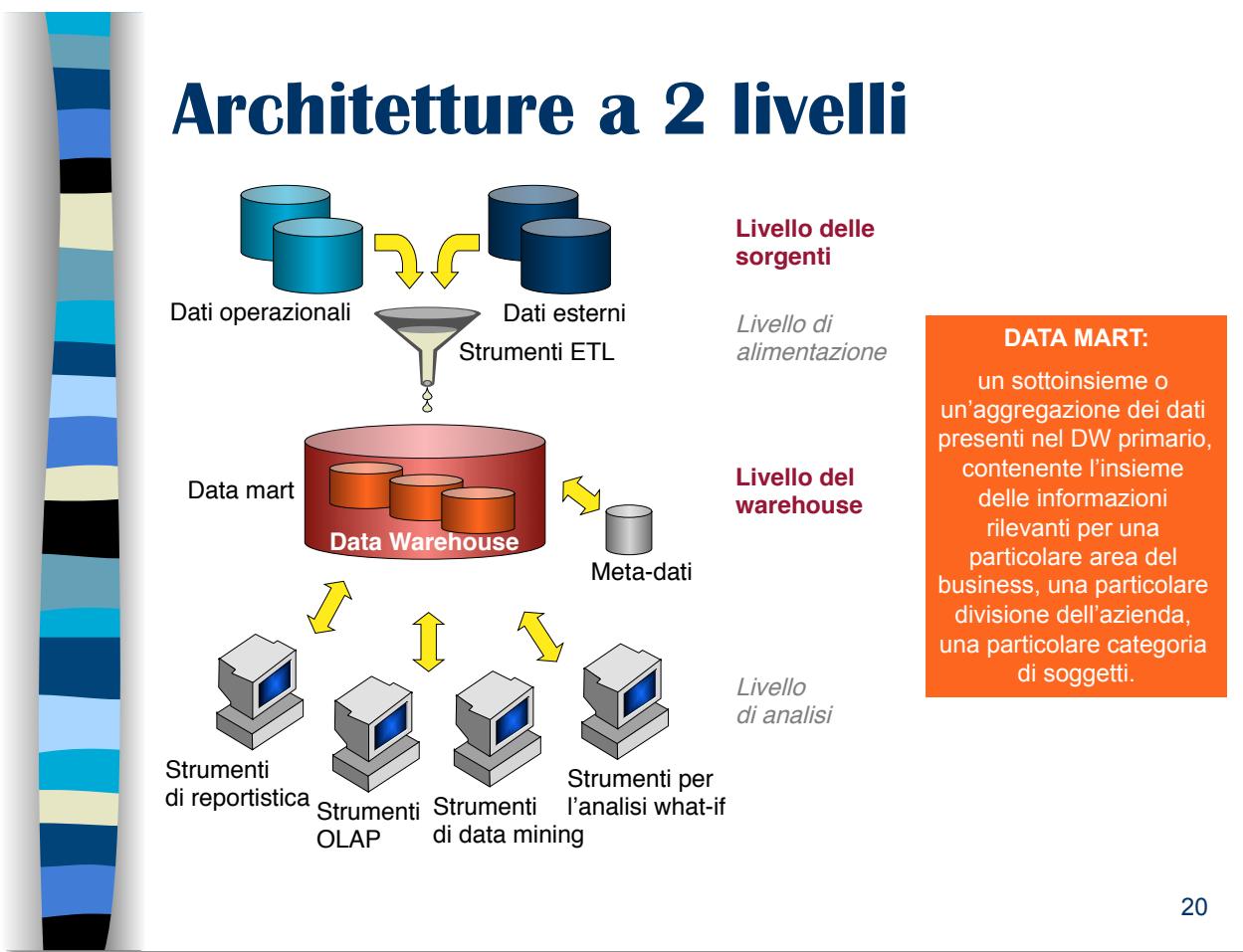
18

Architetture a 1 livello

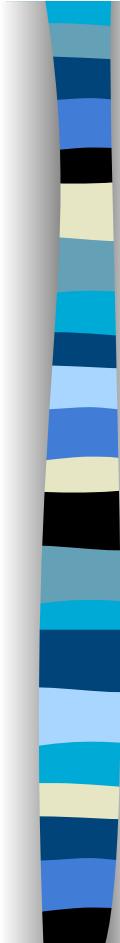


19

Architetture a 2 livelli



20



Architetture a 2 livelli

- I data mart alimentati dal DW primario sono detti *dipendenti*. Per i sistemi collocati all'interno di realtà aziendali medio-grandi essi sono utili:
 - ✓ come blocchi costruttivi durante la realizzazione incrementale del DW;
 - ✓ in quanto delineano i contorni delle informazioni necessarie a un particolare tipo di utenti per le loro interrogazioni;
 - ✓ poiché, essendo di dimensioni inferiori al DW primario, permettono di raggiungere prestazioni migliori
- In alcuni contesti si preferisce adottare data mart alimentati direttamente dalle sorgenti, detti *indipendenti*
 - ✓ L'assenza di un DW primario snellisce le fasi progettuali, ma determina uno schema complesso di accessi ai dati e ingenera il rischio di inconsistenze tra i data mart

21

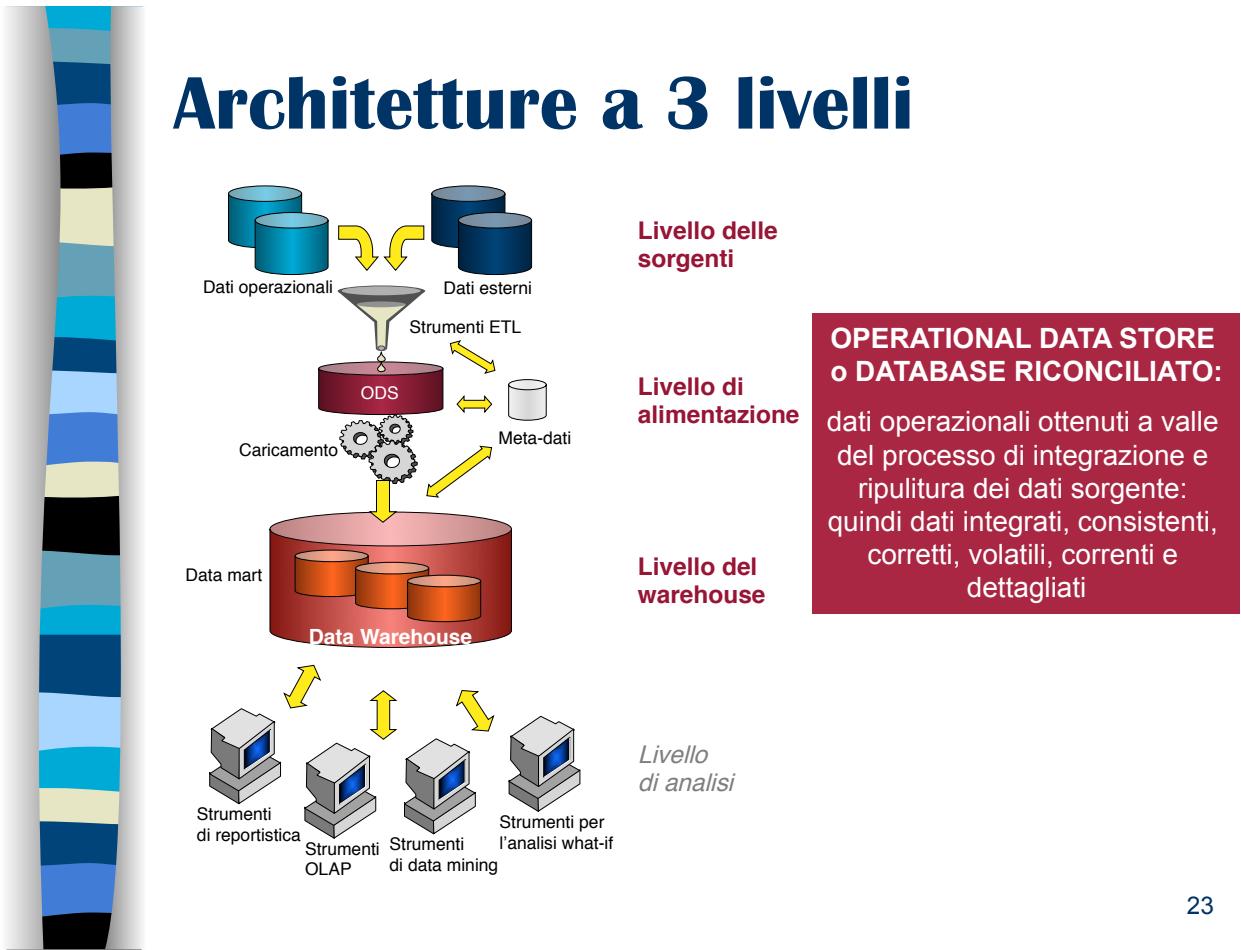


Architetture a 2 livelli

- Vantaggi:
 - ✓ A livello del warehouse è continuamente disponibile informazione di buona qualità anche quando, per motivi tecnici oppure organizzativi, è temporaneamente precluso l'accesso alle sorgenti
 - ✓ L'interrogazione analitica effettuata sul DW non interferisce con la gestione delle transazioni a livello operazionale, la cui affidabilità è essenziale per il funzionamento dell'azienda
 - ✓ L'organizzazione logica del DW è basata sul modello multidimensionale, mentre le sorgenti offrono in genere modelli relazionali o semi-strutturati
 - ✓ C'è una discordanza temporale e di granularità tra sistemi OLTP, che trattano dati correnti e al massimo livello di dettaglio, e sistemi OLAP che operano su dati storici e di sintesi
 - ✓ A livello del warehouse è possibile impiegare tecniche specifiche per ottimizzare le prestazioni per applicazioni di analisi e reportistica

22

Architetture a 3 livelli

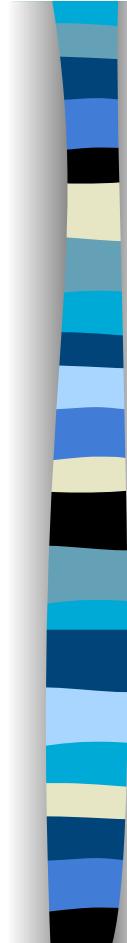


23

Architetture a 3 livelli

- Il vantaggio principale del livello dei dati riconciliati è che esso crea un modello di dati comune e di riferimento per l'intera azienda, introducendo al contempo una separazione netta tra le problematiche legate all'estrazione e integrazione dei dati dalle sorgenti e quelle inerenti l'alimentazione del DW
- D'altro canto, i dati riconciliati introducono un'ulteriore ridondanza rispetto ai dati operazionali sorgente

24

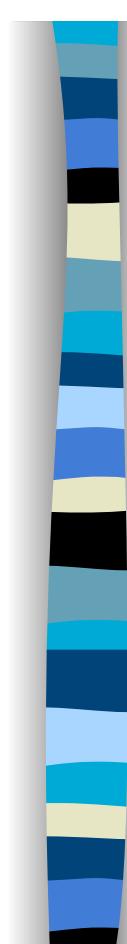


Riconciliazione... terminologica

- Data Warehouse
 - ✓ Contiene dati denormalizzati (tipicamente star e snowflake schema)
- Data Mart
 - ✓ Corrisponde a una porzione del data warehouse
 - ✓ Include più cubi multidimensionali, per cui “data mart” non è sinonimo di “cubo multidimensionale”
- Cubo multidimensionale
 - ✓ Implementa un fatto che lo modella a livello concettuale
- Misura
 - ✓ Chiamata anche “metrica” o “KPI”
- Database riconciliato o Operational Data Store (ODS)
 - ✓ Se presente, è parte integrante dell’architettura
 - ✓ Espone un modello coerente del business, contiene dati normalizzati e può essere usato per la reportistica operativa
 - ✓ Non è sinonimo di “data warehouse” (anche se alcuni lo chiamano erroneamente “data warehouse”)
- Staging area
 - ✓ Area di lavoro dell’ETL, spesso erroneamente confusa con l’ODS
 - ✓ Ha un ruolo di servizio all’interno dell’architettura

© M. Golfarelli e S. Rizzi 2011

25



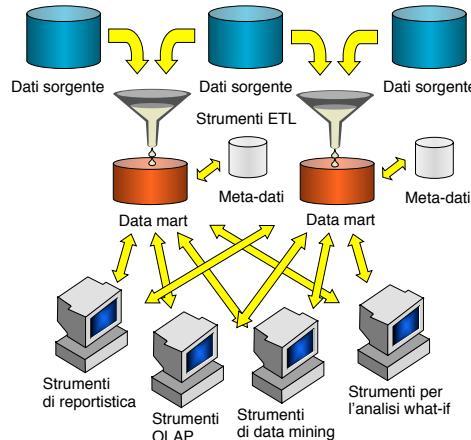
Architetture: un’altra classificazione

- Data mart indipendenti
- Data mart bus
- Hub-and-spoke
- Federazione

26

Data mart indipendenti

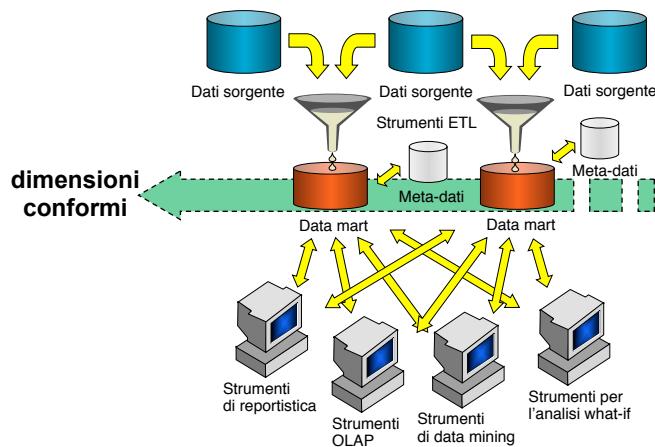
- Primo approccio al data warehousing
- Problema dell'inconsistenza (*data silos*)



27

Data mart bus

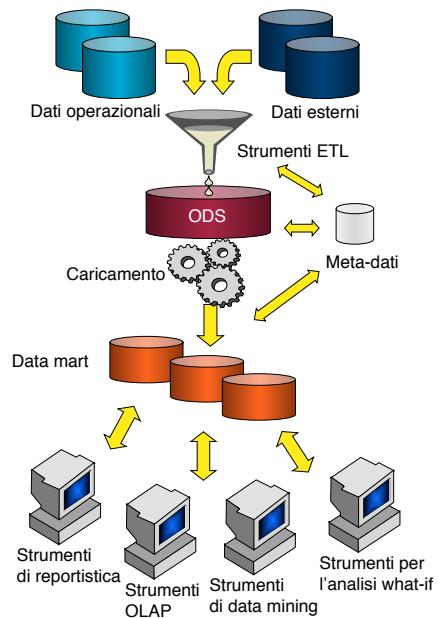
- Approccio consigliato da Kimball
- Integrazione a livello logico
- “Enterprise view”



28

Hub-and-spoke

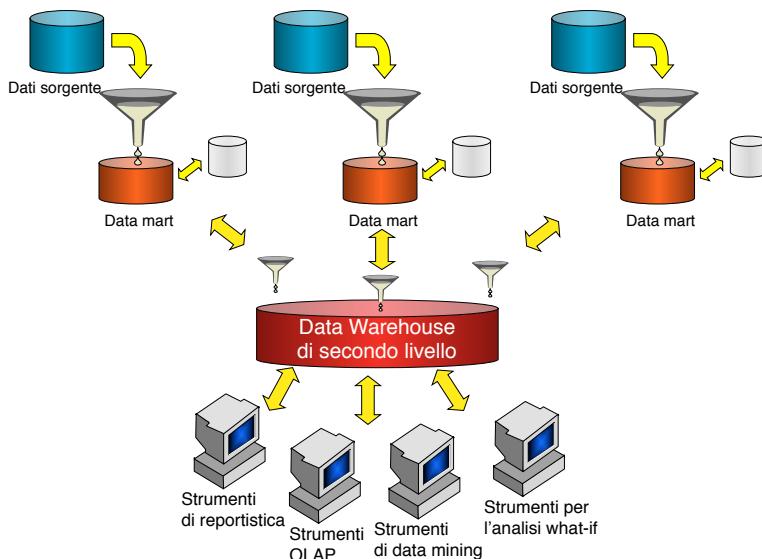
- Una delle architetture più usate in contesti medio-grandi



29

Federazione

- Ideale per contesti molto dinamici (fusioni-acquisizioni)
- Problema dell'integrazione efficace ed efficiente



30



Fattori di scelta dell'architettura

- Interdipendenza informativa tra le unità organizzative aziendali
 - ✓ incoraggia l'adozione di architetture enterprise-wide
- Urgenza del progetto di data warehousing
 - ✓ incoraggia l'adozione di architetture "veloci"
- Vincoli sulle risorse economiche e umane
- Ruolo del progetto di warehousing all'interno della strategia aziendale
 - ✓ data mart indipendenti vs. hub-and-spoke
- Compatibilità con piattaforme esistenti
- Capacità dello staff IT
- Posizione organizzativa dello sponsor di progetto
 - ✓ architetture aziendali vs. architetture dipartimentali

31

ETL

- Il ruolo degli strumenti di *Extraction, Transformation and Loading* è quello di alimentare una sorgente dati singola, dettagliata, esaurente e di alta qualità che possa a sua volta alimentare il DW (*riconciliazione*)
- Durante il processo di alimentazione del DW, la riconciliazione avviene in due occasioni: quando il DW viene popolato per la prima volta, e periodicamente quando il DW viene aggiornato.
 - ✓ estrazione
 - ✓ pulitura
 - ✓ trasformazione
 - ✓ caricamento

32

Estrazione

- I dati rilevanti vengono estratti dalle sorgenti
 - ✓ L'estrazione **statica** viene effettuata quando il DW deve essere popolato per la prima volta e consiste concettualmente in una fotografia dei dati operazionali
 - ✓ L'estrazione **incrementale** viene usata per l'aggiornamento periodico del DW, e cattura solamente i cambiamenti avvenuti nelle sorgenti dall'ultima estrazione
 - basata sul log mantenuto dal DBMS operazionale
 - basata su time-stamp
 - guidata dalle sorgenti
- La scelta dei dati da estrarre avviene principalmente in base alla loro qualità



33

Pulitura

- Si incarica di migliorare la qualità dei dati delle sorgenti
 - ✓ dati duplicati
 - ✓ inconsistenza tra valori logicamente associati
 - ✓ dati mancanti
 - ✓ uso non previsto di un campo
 - ✓ valori impossibili o errati
 - ✓ valori inconsistenti per la stessa entità dovuti a errori di battitura



34

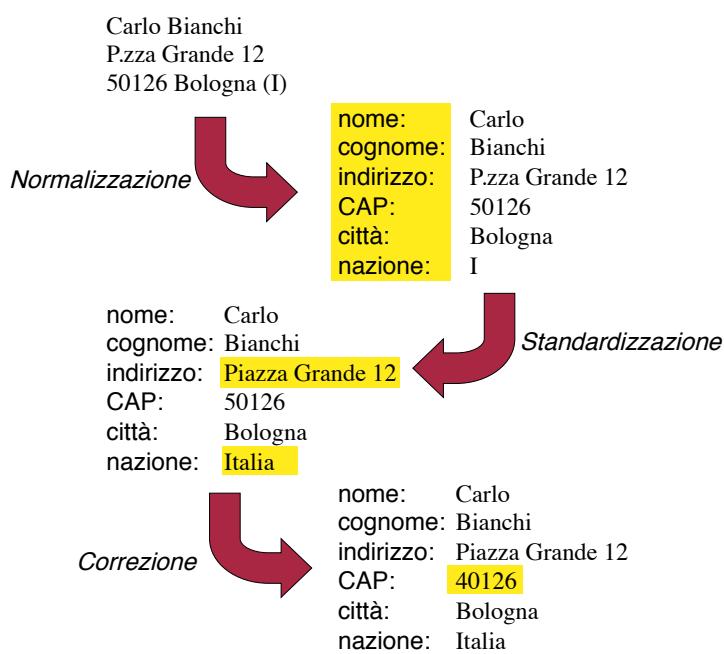
Trasformazione

- Converte i dati dal formato operazionale sorgente a quello del DW. La corrispondenza con il livello sorgente è complicata dalla presenza di fonti distinte eterogenee, che richiede una complessa fase di integrazione
 - ✓ presenza di testi liberi che nascondono informazioni importanti
 - ✓ utilizzo di formati e convenzioni differenti per lo stesso dato
- Per l'alimentazione dei dati riconciliati:
 - ✓ conversione e normalizzazione (operano a livello di formato di memorizzazione e di unità di misura per uniformare i dati)
 - ✓ matching (stabilisce corrispondenze tra campi equivalenti in sorgenti diverse)
 - ✓ selezione (riduce il numero di campi e di record rispetto alle sorgenti)
- Per l'alimentazione del DW:
 - ✓ la normalizzazione è sostituita dalla denormalizzazione
 - ✓ si introduce l'aggregazione, che realizza le opportune sintesi dei dati



35

Pulitura e trasformazione



36

Caricamento

- Il caricamento dei dati nel DW
 - ✓ Refresh: i dati del DW vengono riscritti integralmente, sostituendo quelli precedenti (tecnica utilizzata per popolare inizialmente il DW)
 - ✓ Update: i soli cambiamenti occorsi nei dati sorgente vengono aggiunti nel DW (tecnica utilizzata per l'aggiornamento periodico del DW)



37

Verso il modello multidimensionale

“Che incassi sono stati registrati l’anno passato per ciascuna regione e ciascuna categoria di prodotto?”

“Che rapporto c’è tra l’andamento dei titoli azionari dei produttori di PC e i profitti trimestrali lungo gli ultimi 5 anni?”

“Quali sono le tipologie di ordini che massimizzano gli incassi?”

“Quale di due nuove terapie risulta più efficace ai fini della diminuzione della durata media di un ricovero?”

“Che rapporto c’è tra i profitti realizzati con spedizioni di meno di 10 elementi e quelli realizzati con spedizioni di più di 10 elementi?”

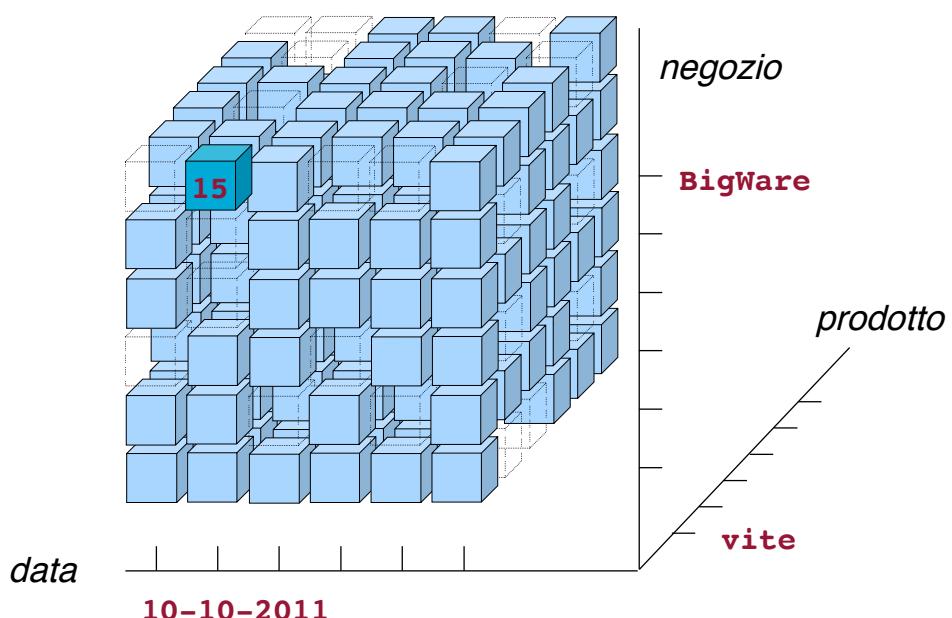
38

Il modello multidimensionale

- È il fondamento per la rappresentazione e l'interrogazione dei dati nei data warehouse.
- I *fatti* di interesse sono rappresentati in *cubi* in cui:
 - ✓ ogni cella contiene *misure* numeriche che quantificano il fatto da diversi punti di vista;
 - ✓ ogni asse rappresenta una *dimensione* di interesse per l'analisi;
 - ✓ ogni dimensione può essere la radice di una *gerarchia* di attributi usati per aggregare i dati memorizzati nei cubi base.

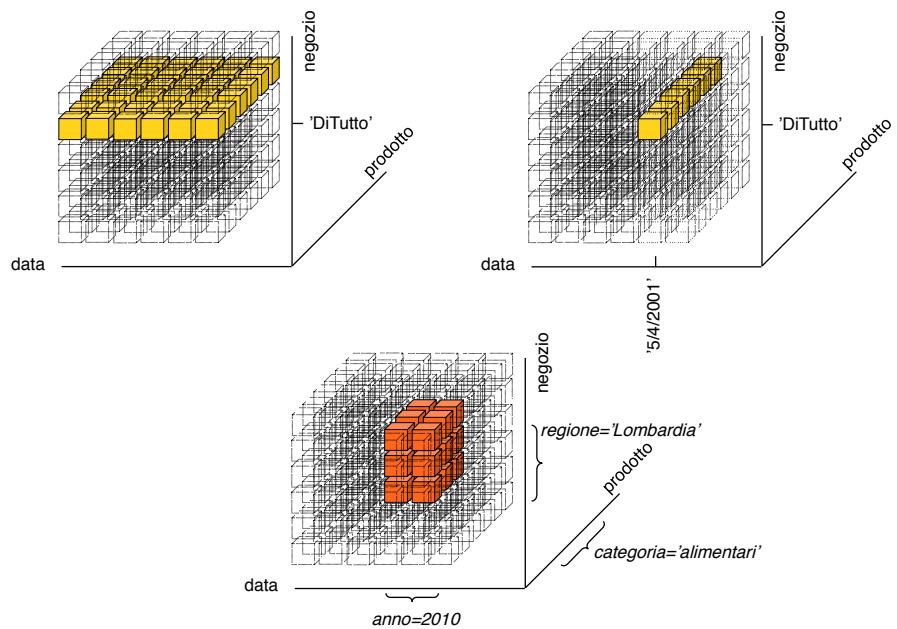
39

Il cubo delle vendite



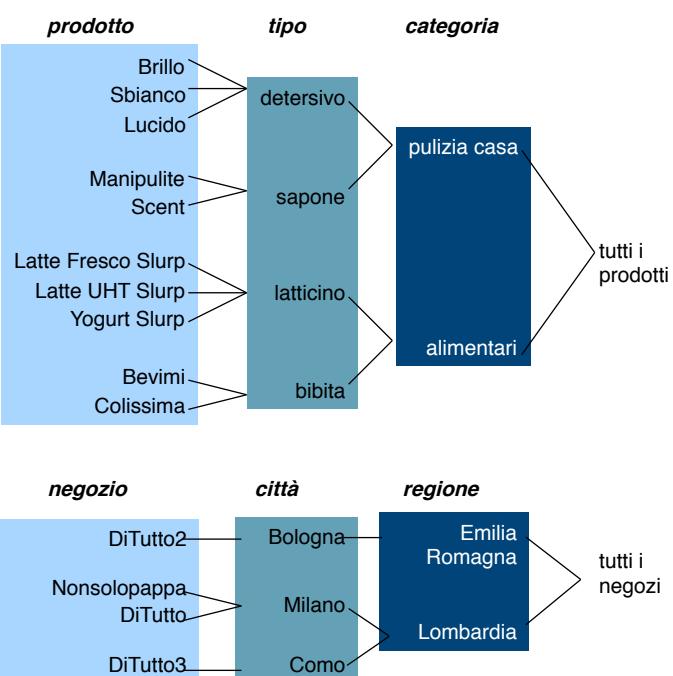
40

Slicing and dicing



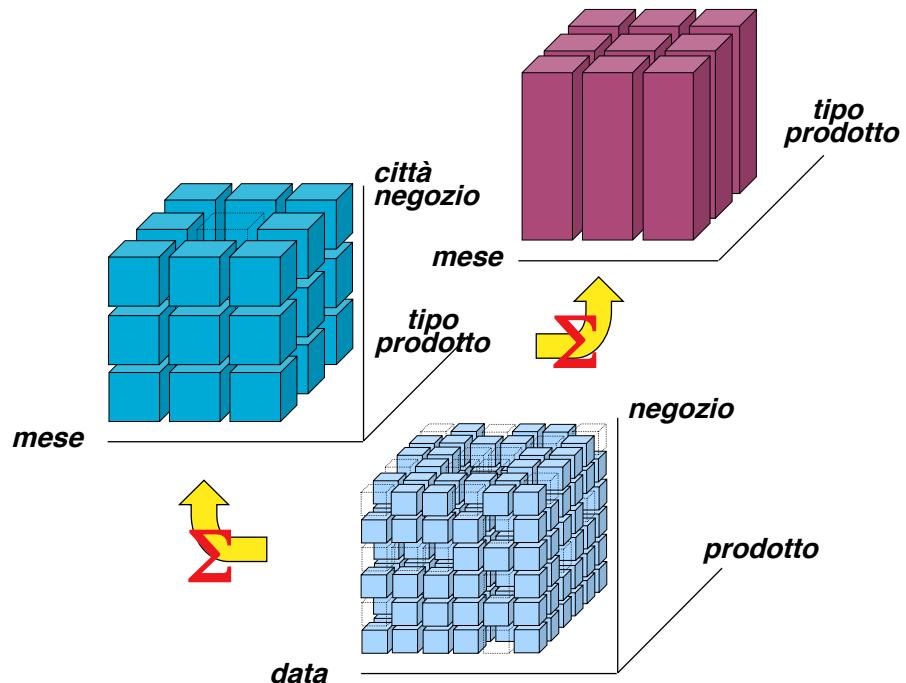
41

Le gerarchie



42

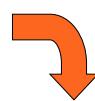
Aggregazione



43

Aggregazione

	DiTutto	DiTutto2	Nonsolopappa
1/1/2000	—	—	—
2/1/2000	10	15	5
3/1/2000	20	—	5
.....
1/1/2001	—	—	—
2/1/2001	15	10	20
3/1/2001	20	20	25
.....
1/1/2002	—	—	—
2/1/2002	20	8	25
3/1/2002	20	12	20
.....



	DiTutto	DiTutto2	Nonsolopappa
Gennaio 2000	200	180	150
Febbraio 2000	180	150	120
Marzo 2000	220	180	160
.....
Gennaio 2001	350	220	200
Febbraio 2001	300	200	250
Marzo 2001	310	180	300
.....
Gennaio 2002	380	200	220
Febbraio 2002	310	200	250
Marzo 2002	300	160	280
.....



	DiTutto	DiTutto2	Nonsolopappa
2000	2400	2000	1600
2001	3200	2300	3000
2002	3400	2200	3200

Total:	DiTutto	DiTutto2	Nonsolopappa
	9000	6500	7800

44

Tecniche di analisi dei dati

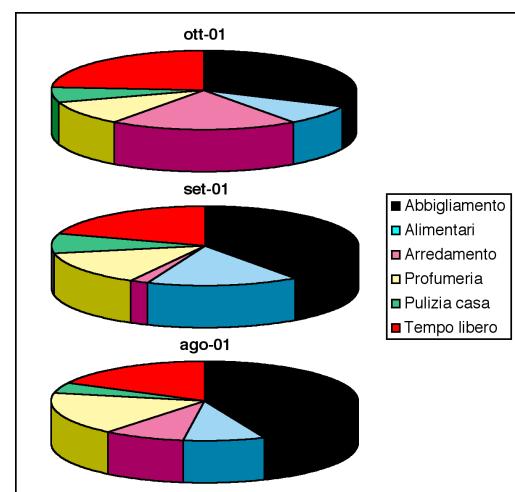
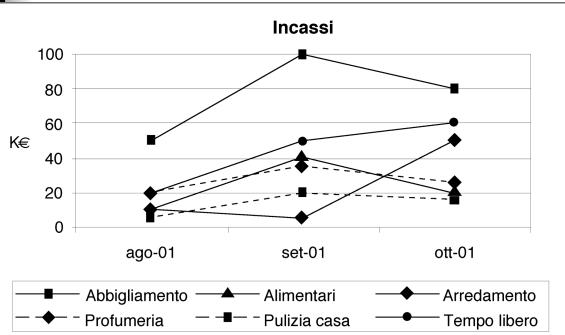
- Una volta che i dati sono stati ripuliti, integrati e trasformati, occorre capire come trarne il massimo vantaggio informativo
- Esistono due approcci differenti, supportati da altrettante categorie di strumenti, all'interrogazione di un DW da parte degli utenti finali:
 - ✓ *reportistica*: non richiede conoscenze informatiche
 - ✓ *OLAP*: richiede all'utente di ragionare in modo multidimensionale e di conoscere l'interfaccia dello strumento grafico utilizzato

45

Reportistica

orientato agli utenti
che hanno necessità
di accedere, a
intervalli di tempo
predefiniti, a
informazioni
strutturate in modo
pressoché invariabile

incassi (K€)	Ottobre 2001	Settembre 2001	Agosto 2001
Abbigliamento	80	100	50
Alimentari	20	40	10
Arredamento	50	5	10
Profumeria	25	35	20
Pulizia casa	15	20	5
Tempo libero	60	50	20



46

Reportistica



47

OLAP

- È la principale modalità di fruizione delle informazioni contenute in un DW
- Consente, a utenti le cui necessità di analisi non siano facilmente identificabili a priori, di analizzare ed esplorare interattivamente i dati sulla base del modello multidimensionale
- Mentre gli utenti degli strumenti di reportistica svolgono un ruolo essenzialmente passivo, gli utenti OLAP sono in grado di costruire attivamente una sessione di analisi complessa in cui ciascun passo effettuato è conseguenza dei risultati ottenuti al passo precedente
 - ✓ estemporaneità delle sessioni di lavoro
 - ✓ richiesta approfondita conoscenza dei dati
 - ✓ complessità delle interrogazioni formulabili
 - ✓ orientamento verso utenti non esperti di informatica



interfaccia flessibile, facile
da usare ed efficace

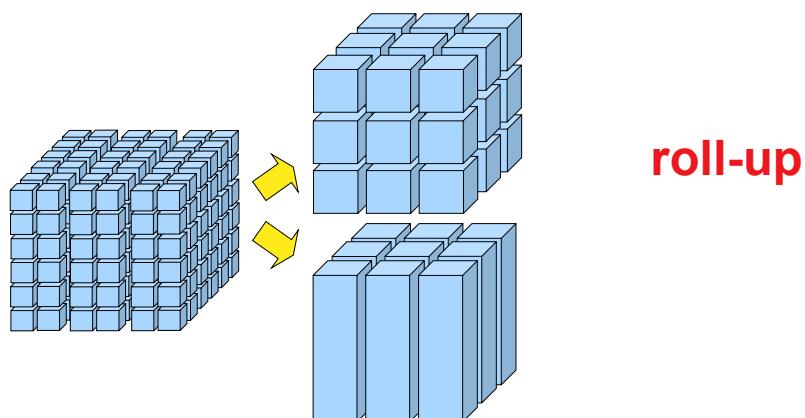
48

OLAP: sessione

- Una sessione OLAP consiste in un *percorso di navigazione* che riflette il procedimento di analisi di uno o più fatti di interesse sotto diversi aspetti e a diversi livelli di dettaglio. Questo percorso si concretizza in una sequenza di interrogazioni spesso formulate non direttamente, ma per differenza rispetto all'interrogazione precedente
- Ogni passo della sessione di analisi è scandito dall'applicazione di un **operatore OLAP** che trasforma l'ultima interrogazione formulata in una nuova interrogazione
- Il risultato delle interrogazioni è di tipo multidimensionale; gli strumenti OLAP rappresentano tipicamente i dati in modo tabellare evidenziando le diverse dimensioni mediante intestazioni multiple, colori ecc.

49

OLAP: operatori



50

OLAP: operatori

	Metrics Customer Region	Dollar Sales										
Month		North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany	Canada
Jan 97		\$ 620	\$ 753	\$ 30	\$ 660	\$ 2,405	\$ 1,312	\$ 440	\$ 1,002	\$ 1,002	\$ 383	\$ 210
Feb 97		\$ 258	\$ 252	\$ 800	\$ 975	\$ 160	\$ 582	\$ 744	\$ 310	\$ 799	\$ 118	\$ 357
Mar 97		\$ 648	\$ 244	\$ 148	\$ 250	\$ 1,085	\$ 2,961	\$ 650	\$ 1,240	\$ 119	\$ 142	\$ 96
Apr 97		\$ 787	\$ 588	\$ 447	\$ 486	\$ 226	\$ 506	\$ 601	\$ 119	\$ 550	\$ 85	
May 97		\$ 1,350	\$ 245	\$ 936	\$ 159	\$ 664	\$ 626	\$ 107	\$ 135	\$ 200	\$ 177	\$ 230
Jun 97		\$ 842	\$ 582	\$ 1,281	\$ 937	\$ 240	\$ 774	\$ 176	\$ 1,139	\$ 652	\$ 254	\$ 745
Jul 97		\$ 652	\$ 690	\$ 486	\$ 1,293	\$ 605	\$ 303	\$ 818	\$ 103	\$ 124	\$ 173	\$ 66
Aug 97		\$ 1,783	\$ 304	\$ 1,032	\$ 170	\$ 398	\$ 356	\$ 432	\$ 190	\$ 241	\$ 407	\$ 259
Sep 97		\$ 581	\$ 778	\$ 3,558	\$ 587	\$ 440	\$ 1,652	\$ 1,071	\$ 315	\$ 210	\$ 202	
Oct 97		\$ 2,291	\$ 1,840	\$ 600	\$ 656	\$ 1,300	\$ 718	\$ 1,210	\$ 427	\$ 220	\$ 520	\$ 65
Nov 97		\$ 39	\$ 1,602	\$ 1,082	\$ 1,187	\$ 842	\$ 759	\$ 745	\$ 232	\$ 101	\$ 1,037	\$ 37
Dec 97		\$ 381	\$ 1,586	\$ 343	\$ 118	\$ 1,459	\$ 635	\$ 2,021	\$ 259	\$ 210	\$ 119	\$ 189
Jan 98		\$ 311	\$ 1,174	\$ 2,634	\$ 3,130	\$ 954	\$ 2,083	\$ 1,351	\$ 747	\$ 426	\$ 447	\$ 1,141
Feb 98		\$ 2,518	\$ 702	\$ 1,123	\$ 1,336	\$ 1,227	\$ 3,887	\$ 545	\$ 268	\$ 277	\$ 282	
Mar 98		\$ 2,459	\$ 1,523	\$ 1,178	\$ 4,708	\$ 1,420	\$ 1,948	\$ 1,705	\$ 276	\$ 1,168	\$ 63	
Apr 98		\$ 407	\$ 841	\$ 524	\$ 712	\$ 133	\$ 2,486	\$ 49	\$ 390	\$ 1,298	\$ 221	\$ 46
May 98		\$ 667	\$ 1,721	\$ 440	\$ 148	\$ 80	\$ 1,310	\$ 303	\$ 104	\$ 657	\$ 65	
Jun 98		\$ 699	\$ 1,096	\$ 898	\$ 353	\$ 902	\$ 839	\$ 230	\$ 155	\$ 105	\$ 75	
Jul 98		\$ 586	\$ 1,897	\$ 412	\$ 226	\$ 406	\$ 361	\$ 1,628	\$ 267	\$ 1,011	\$ 41	\$ 184
Aug 98		\$ 894	\$ 326	\$ 792	\$ 1,832	\$ 1,199	\$ 295	\$ 1,816	\$ 277	\$ 102	\$ 118	\$ 115
Sep 98		\$ 338	\$ 3,179	\$ 505	\$ 427	\$ 99	\$ 2,976	\$ 885	\$ 135	\$ 85	\$ 1,110	\$ 510
Oct 98		\$ 544	\$ 413	\$ 1,467	\$ 209	\$ 679	\$ 706	\$ 556	\$ 480	\$ 485	\$ 99	\$ 160
Nov 98		\$ 671	\$ 459	\$ 1,471	\$ 2,066	\$ 701	\$ 716	\$ 986	\$ 1,227	\$ 154	\$ 440	\$ 361
Dec 98		\$ 836	\$ 2,096	\$ 1,726	\$ 3,642	\$ 395	\$ 1,740	\$ 1,943	\$ 1,143	\$ 366	\$ 307	\$ 118

roll-up

	Metrics Customer Region	Dollar Sales										
Quarter		North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany	Canada
Q1 1997		\$ 1,526	\$ 1,249	\$ 978	\$ 1,895	\$ 3,650	\$ 4,855	\$ 1,834	\$ 2,552	\$ 1,920	\$ 643	\$ 663
Q2 1997		\$ 2,979	\$ 1,415	\$ 2,664	\$ 1,582	\$ 1,130	\$ 1,906	\$ 884	\$ 1,393	\$ 1,402	\$ 516	\$ 975
Q3 1997		\$ 3,016	\$ 1,772	\$ 5,076	\$ 2,050	\$ 1,443	\$ 2,311	\$ 2,321	\$ 608	\$ 575	\$ 782	\$ 325
Q4 1997		\$ 2,711	\$ 5,030	\$ 2,025	\$ 1,961	\$ 3,601	\$ 2,112	\$ 3,976	\$ 918	\$ 531	\$ 1,676	\$ 291
Q1 1998		\$ 5,288	\$ 3,399	\$ 4,935	\$ 9,174	\$ 3,601	\$ 9,484	\$ 3,844	\$ 2,720	\$ 979	\$ 1,897	\$ 1,204
Q2 1998		\$ 1,772	\$ 3,658	\$ 1,862	\$ 2,123	\$ 1,115	\$ 4,635	\$ 352	\$ 724	\$ 2,110	\$ 391	\$ 121
Q3 1998		\$ 1,818	\$ 5,402	\$ 1,709	\$ 2,485	\$ 1,704	\$ 3,632	\$ 4,329	\$ 679	\$ 1,198	\$ 1,269	\$ 809
Q4 1998		\$ 2,051	\$ 2,968	\$ 4,664	\$ 5,917	\$ 1,775	\$ 3,162	\$ 3,485	\$ 2,750	\$ 1,005	\$ 846	\$ 639

51

OLAP: operatori

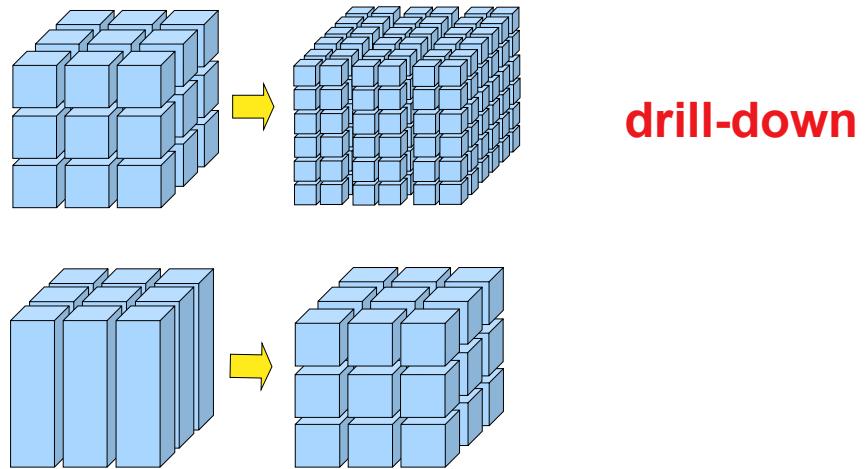
	Metrics Customer Region	Dollar Sales										
Category	Year	North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany	Canada
Electronics	1997	\$ 138	\$ 1,774	\$ 384	\$ 138	\$ 2,346	\$ 2,554	\$ 2,184	\$ 566	\$ 199	\$	
	1998	\$ 1,184	\$ 4,529	\$ 1,892	\$ 7,232	\$ 651	\$ 9,488	\$ 476	\$ 2,683	\$ 462	\$ 7	
Food	1997	\$ 759	\$ 682	\$ 729	\$ 262	\$ 588	\$ 469	\$ 807	\$ 156	\$ 615	\$ 1	
	1998	\$ 538	\$ 925	\$ 959	\$ 677	\$ 213	\$ 1,503	\$ 261	\$ 165	\$ 175	\$ 1	
Gifts	1997	\$ 2,532	\$ 1,355	\$ 1,854	\$ 1,413	\$ 2,535	\$ 2,132	\$ 1,904	\$ 908	\$ 375	\$ 1,0	
	1998	\$ 1,955	\$ 2,785	\$ 2,800	\$ 2,695	\$ 1,813	\$ 2,844	\$ 1,778	\$ 1,158	\$ 717	\$ 6	
Health & Beauty	1997	\$ 624	\$ 640	\$ 1,317	\$ 647	\$ 588	\$ 754	\$ 654	\$ 143	\$ 292	\$ 3	
	1998	\$ 611	\$ 887	\$ 566	\$ 382	\$ 499	\$ 1,162	\$ 1,044	\$ 273	\$ 72		
Household	1997	\$ 5,354	\$ 4,112	\$ 5,410	\$ 4,446	\$ 3,058	\$ 3,974	\$ 2,654	\$ 3,545	\$ 2,875	\$ 1,9	
	1998	\$ 5,787	\$ 5,320	\$ 5,416	\$ 6,812	\$ 4,334	\$ 5,008	\$ 7,588	\$ 2,139	\$ 3,649	\$ 2,7	
Kid's Korner	1997	\$ 201	\$ 398	\$ 485	\$ 186	\$ 409	\$ 323	\$ 396	\$ 105	\$ 34	\$	
	1998	\$ 247	\$ 422	\$ 441	\$ 380	\$ 221	\$ 592	\$ 290	\$ 198	\$ 19	\$	
Travel	1997	\$ 624	\$ 505	\$ 564	\$ 386	\$ 300	\$ 978	\$ 416	\$ 48	\$ 38		
	1998	\$ 608	\$ 559	\$ 1,096	\$ 611	\$ 454	\$ 316	\$ 573	\$ 257	\$ 198	\$	

roll-up

	Metrics	Dollar Sales
Category	Year	
Electronics	1997	\$ 10,616
	1998	\$ 29,299
Food	1997	\$ 5,300
	1998	\$ 5,638
Gifts	1997	\$ 16,315
	1998	\$ 20,047
Health & Beauty	1997	\$ 6,042
	1998	\$ 5,665
Household	1997	\$ 38,383
	1998	\$ 50,391
Kid's Korner	1997	\$ 2,559
	1998	\$ 2,943
Travel	1997	\$ 4,497
	1998	\$ 4,792

52

OLAP: operatori



53

OLAP: operatori

	Metrics	Dollar Sales	North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany	Canada
Quarter	Customer Region												
Q1 1997		\$ 1.526	\$ 1.249	\$ 978	\$ 1.885	\$ 3.650	\$ 4.855	\$ 1.834	\$ 2.552	\$ 1.920	\$ 643	\$ 663	
Q2 1997		\$ 2.979	\$ 1.415	\$ 2.664	\$ 1.582	\$ 1.130	\$ 1.906	\$ 884	\$ 1.393	\$ 1.402	\$ 516	\$ 975	
Q3 1997		\$ 3.016	\$ 1.772	\$ 5.076	\$ 2.050	\$ 1.443	\$ 2.311	\$ 2.321	\$ 608	\$ 575	\$ 782	\$ 325	
Q4 1997		\$ 2.711	\$ 5.030	\$ 2.025	\$ 1.961	\$ 3.601	\$ 2.112	\$ 3.976	\$ 918	\$ 531	\$ 1.676	\$ 291	
Q1 1998		\$ 5.288	\$ 3.399	\$ 4.935	\$ 9.174	\$ 3.601	\$ 9.484	\$ 3.844	\$ 2.720	\$ 979	\$ 1.897	\$ 1.201	
Q2 1998		\$ 1.773	\$ 3.658	\$ 1.862	\$ 1.213	\$ 1.115	\$ 4.635	\$ 352	\$ 724	\$ 2.110	\$ 391	\$ 12	
Q3 1998		\$ 1.818	\$ 5.402	\$ 1.709	\$ 2.485	\$ 1.704	\$ 3.632	\$ 4.329	\$ 679	\$ 1.198	\$ 1.269	\$ 809	
Q4 1998		\$ 2.051	\$ 2.968	\$ 4.664	\$ 5.917	\$ 1.775	\$ 3.162	\$ 3.485	\$ 2.750	\$ 1.005	\$ 846	\$ 639	

drill-down



	Metrics	Dollar Sales	Arlin	San Pedro	Springfield	Chappel Hill	Scranburg	Pebble Beach	Martinsville	Maddon	Peoria	Pecos	Lake Barkley	Alameda	Fingers Lake
Quarter	Customer City														
Q1 1997		\$ 675											\$ 39		
Q2 1997														\$ 135	
Q3 1997														\$ 252	\$ 63
Q4 1997		\$ 215	\$ 124											\$ 79	\$ 98
Q1 1998														\$ 237	\$ 30
Q2 1998														\$ 30	\$ 119
Q3 1998		\$ 734													
Q4 1998															

54

OLAP: operatori

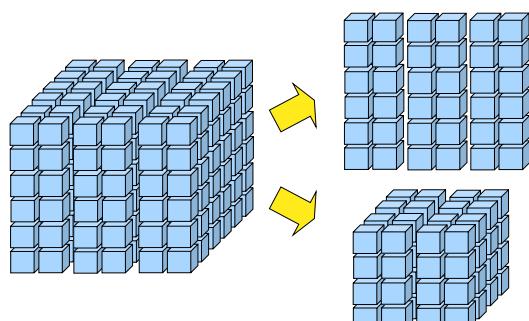
Category	Metrics		Dollar Sales		1997	1998	
	Year	1997	1998	1997	1998		
Electronics	\$ 10,616	\$ 29,299					
Food	\$ 5,900	\$ 5,638					
Gifts	\$ 16,315	\$ 20,047					
Health & Beauty	\$ 6,042	\$ 5,665					
Household	\$ 38,383	\$ 50,391					
Kid's Korner	\$ 2,559	\$ 2,943					
Travel	\$ 4,497	\$ 4,792					

drill-down

Category	Metrics		Dollar Sales		North-East		Mid-Atlantic		South-East		Central		South		North-West	
	Customer Region	Year	North-East		Mid-Atlantic		South-East		Central		South		North-West		Europe	
			1997	1998	1997	1998	1997	1998	1997	1998	1997	1998	1997	1998	1997	1998
Electronics	\$ 138	\$ 1,184	\$ 1,774	\$ 4,529	\$ 384	\$ 1,892	\$ 138	\$ 7,232	\$ 2,346	\$ 651	\$ 2,554	\$ 9,488	\$ 1,184	\$ 1,774	\$ 4,529	\$ 384
Food	\$ 759	\$ 538	\$ 682	\$ 925	\$ 729	\$ 959	\$ 262	\$ 677	\$ 588	\$ 213	\$ 469	\$ 1,503	\$ 682	\$ 959	\$ 262	\$ 677
Gifts	\$ 2,532	\$ 1,955	\$ 1,355	\$ 2,785	\$ 1,854	\$ 2,800	\$ 1,413	\$ 2,695	\$ 2,535	\$ 1,813	\$ 2,132	\$ 2,844	\$ 1,355	\$ 2,785	\$ 1,854	\$ 2,800
Health & Beauty	\$ 624	\$ 611	\$ 640	\$ 887	\$ 1,317	\$ 566	\$ 647	\$ 382	\$ 588	\$ 499	\$ 754	\$ 1,162	\$ 566	\$ 647	\$ 382	\$ 588
Household	\$ 5,354	\$ 5,787	\$ 4,112	\$ 5,320	\$ 5,410	\$ 5,416	\$ 4,446	\$ 6,012	\$ 3,058	\$ 4,334	\$ 3,974	\$ 5,008	\$ 4,446	\$ 6,012	\$ 3,058	\$ 4,334
Kid's Korner	\$ 201	\$ 247	\$ 398	\$ 422	\$ 485	\$ 441	\$ 186	\$ 380	\$ 409	\$ 221	\$ 323	\$ 592	\$ 441	\$ 186	\$ 380	\$ 409
Travel	\$ 624	\$ 608	\$ 505	\$ 559	\$ 564	\$ 1,096	\$ 386	\$ 611	\$ 300	\$ 464	\$ 978	\$ 316	\$ 564	\$ 1,096	\$ 386	\$ 611

55

OLAP: operatori



slice-and-dice

56

OLAP: operatori

slice-and-dice

Category	Year	Metrics Customer Region									
		North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany
Electronics	1997	\$ 138	\$ 1.774	\$ 384	\$ 138	\$ 2.346	\$ 2.554	\$ 2.184	\$ 566	\$ 199	\$ 7
Food	1997	\$ 759	\$ 682	\$ 729	\$ 262	\$ 588	\$ 469	\$ 807	\$ 156	\$ 615	\$ 1
Gifts	1997	\$ 538	\$ 925	\$ 959	\$ 677	\$ 213	\$ 1.503	\$ 261	\$ 165	\$ 175	\$ 1
Health & Beauty	1997	\$ 2.532	\$ 1.355	\$ 1.854	\$ 1.413	\$ 2.535	\$ 2.132	\$ 1.904	\$ 908	\$ 375	\$ 1.0
Household	1997	\$ 624	\$ 640	\$ 1.317	\$ 647	\$ 588	\$ 754	\$ 654	\$ 143	\$ 292	\$ 3
Kid's Korner	1997	\$ 611	\$ 887	\$ 566	\$ 382	\$ 499	\$ 1.162	\$ 1.044	\$ 273	\$ 72	
Travel	1997	\$ 5.787	\$ 5.320	\$ 5.416	\$ 6.812	\$ 4.334	\$ 5.008	\$ 7.588	\$ 2.139	\$ 3.649	\$ 1.9
	1998	\$ 608	\$ 559	\$ 1.096	\$ 611	\$ 464	\$ 316	\$ 573	\$ 257	\$ 198	

Filter Details: Year = 1998											
Category	Customer Region	Metrics									
		North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany
Electronics		\$ 1.184	\$ 4.529	\$ 1.892	\$ 7.232	\$ 651	\$ 9.488	\$ 476	\$ 2.683	\$ 462	\$ 702
Food		\$ 538	\$ 925	\$ 959	\$ 677	\$ 213	\$ 1.503	\$ 261	\$ 165	\$ 175	\$ 100
Gifts		\$ 1.955	\$ 2.785	\$ 2.800	\$ 2.695	\$ 1.813	\$ 2.844	\$ 1.778	\$ 1.158	\$ 717	\$ 686
Health & Beauty		\$ 611	\$ 887	\$ 566	\$ 382	\$ 499	\$ 1.162	\$ 1.044	\$ 273	\$ 72	
Household		\$ 5.787	\$ 5.320	\$ 5.416	\$ 6.812	\$ 4.334	\$ 5.008	\$ 7.588	\$ 2.139	\$ 3.649	\$ 1.9
Kid's Korner		\$ 247	\$ 422	\$ 441	\$ 380	\$ 221	\$ 592	\$ 290	\$ 198	\$ 19	\$ 69
Travel		\$ 624	\$ 505	\$ 564	\$ 386	\$ 300	\$ 978	\$ 416	\$ 48	\$ 38	
		\$ 608	\$ 559	\$ 1.096	\$ 611	\$ 464	\$ 316	\$ 573	\$ 257	\$ 198	\$ 55

57

OLAP: operatori

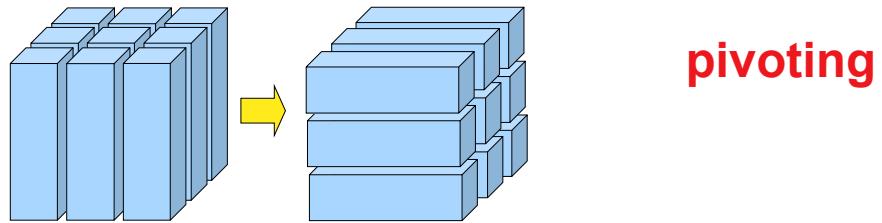
slice-and-dice

Subcategory	Customer City	Metrics										
		Afton	Akron	Albon	Alcameda	Alka	Allagash	Alta	Altoola	Amestra	Amsterdam	Andersonville
Audio						\$ 85						
Automotive	\$ 42	\$ 42		\$ 50			\$ 30					
Chocolate	\$ 30					\$ 20	\$ 22	\$ 44				
Christmas						\$ 25	\$ 30	\$ 15				
Classic Toys						\$ 7	\$ 26					\$ 38
Coffee						\$ 59						
Comfort							\$ 485					
Furniture							\$ 199	\$ 79	\$ 79			
Gadgets							\$ 17	\$ 45				\$ 45
Games & Puzzles												
Gift Baskets												
Golf	\$ 25			\$ 55	\$ 43							
Hearth												
Jewelry	\$ 75			\$ 189		\$ 24	\$ 77	\$ 189	\$ 24			
Kitchen						\$ 55	\$ 21					
Lawn & Garden	\$ 75		\$ 100		\$ 15	\$ 63	\$ 100		\$ 180	\$ 67	\$ 40	\$ 1
Learning	\$ 16								\$ 37			
Meat & Cheese												\$ 25
Miscellaneous												
Natural Remedies	\$ 13											
Pets	\$ 215											
Plants & Flowers	\$ 65	\$ 65	\$ 65									
Safety & Security												
Skin Care												
Sleeping												
Toys & Accessories												

Filter Details: Category = Electronics Alka = Dollar Sales > 80 AND Customer Region = North-West AND Year = 1997											
Subcategory	Customer City	Metrics									
		Alta	Armstrong	Avery Heights	Lane	Mt. Everest	San Fransisco				
Audio						\$ 98					
Comfort							\$ 118				
Gadgets								\$ 1.495			
									\$ 199		

58

OLAP: operatori



59

OLAP: operatori

Category	Metrics	Dollar Sales	
		1997	1998
Electronics	1997	\$ 10.616	
	1998	\$ 29.299	
Food	1997	\$ 5.300	
	1998	\$ 5.638	
Gifts	1997	\$ 16.315	
	1998	\$ 20.047	
Health & Beauty	1997	\$ 6.042	
	1998	\$ 5.665	
Household	1997	\$ 38.383	
	1998	\$ 50.391	
Kid's Korner	1997	\$ 2.559	
	1998	\$ 2.943	
Travel	1997	\$ 4.497	
	1998	\$ 4.792	

pivoting

Category	Metrics	Dollar Sales	
		1997	1998
Electronics		\$ 10.616	\$ 29.299
Food		\$ 5.300	\$ 5.638
Gifts		\$ 16.315	\$ 20.047
Health & Beauty		\$ 6.042	\$ 5.665
Household		\$ 38.383	\$ 50.391
Kid's Korner		\$ 2.559	\$ 2.943
Travel		\$ 4.497	\$ 4.792

60

OLAP: operatori

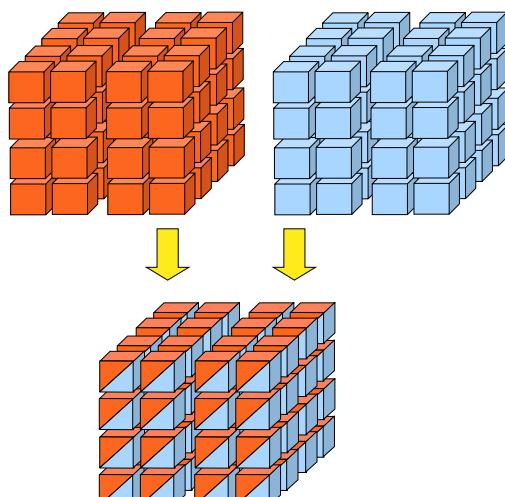
Category	Year	Metrics									
		Customer Region									
		Dollar Sales									
North-East	Mid-Atlantic	South-East	Central	South	North-West	South-West	England	France	Germany		
Electronics	1997	\$ 138	\$ 1.774	\$ 384	\$ 139	\$ 2.346	\$ 2.554	\$ 2.184	\$ 566	\$ 199	\$
	1998	\$ 1.184	\$ 4.529	\$ 1.892	\$ 7.232	\$ 651	\$ 9.488	\$ 476	\$ 2.683	\$ 462	\$ 7
Food	1997	\$ 759	\$ 682	\$ 729	\$ 262	\$ 588	\$ 469	\$ 807	\$ 156	\$ 615	\$ 1
	1998	\$ 538	\$ 925	\$ 959	\$ 677	\$ 213	\$ 1.503	\$ 261	\$ 165	\$ 175	\$ 1
Gifts	1997	\$ 2.532	\$ 1.355	\$ 1.854	\$ 1.413	\$ 2.535	\$ 2.132	\$ 1.904	\$ 908	\$ 375	\$ 1.0
	1998	\$ 1.955	\$ 2.785	\$ 2.800	\$ 2.695	\$ 1.813	\$ 2.844	\$ 1.778	\$ 1.158	\$ 717	\$ 6
Health & Beauty	1997	\$ 624	\$ 640	\$ 1.317	\$ 647	\$ 588	\$ 754	\$ 654	\$ 143	\$ 292	\$ 3
	1998	\$ 611	\$ 887	\$ 566	\$ 382	\$ 499	\$ 1.162	\$ 1.044	\$ 273	\$ 72	
Household	1997	\$ 5.354	\$ 4.112	\$ 5.410	\$ 4.446	\$ 3.058	\$ 3.974	\$ 2.654	\$ 3.545	\$ 2.875	\$ 1.9
	1998	\$ 5.787	\$ 5.320	\$ 5.416	\$ 6.812	\$ 4.334	\$ 5.008	\$ 7.588	\$ 2.139	\$ 3.649	\$ 2.7
Kid's Korner	1997	\$ 201	\$ 398	\$ 485	\$ 186	\$ 409	\$ 323	\$ 396	\$ 105	\$ 34	\$
	1998	\$ 247	\$ 422	\$ 441	\$ 380	\$ 221	\$ 592	\$ 290	\$ 198	\$ 19	\$
Travel	1997	\$ 624	\$ 505	\$ 564	\$ 386	\$ 300	\$ 978	\$ 416	\$ 48	\$ 38	
	1998	\$ 608	\$ 559	\$ 1.096	\$ 611	\$ 464	\$ 316	\$ 573	\$ 257	\$ 198	\$

pivoting

Category	Year	Metrics									
		Customer Region									
		North-East		Mid-Atlantic		South-East		Central		South	
North-East	1997	1997	1998	1997	1998	1997	1998	1997	1998	1997	1998
Electronics	\$ 138	\$ 1.184	\$ 1.774	\$ 4.529	\$ 384	\$ 1.892	\$ 138	\$ 7.232	\$ 2.346	\$ 651	\$ 2.554
Food	\$ 759	\$ 538	\$ 682	\$ 925	\$ 729	\$ 959	\$ 262	\$ 677	\$ 588	\$ 213	\$ 1.503
Gifts	\$ 2.532	\$ 1.955	\$ 1.355	\$ 2.785	\$ 1.854	\$ 2.800	\$ 1.413	\$ 2.695	\$ 2.535	\$ 1.813	\$ 2.132
Health & Beauty	\$ 624	\$ 611	\$ 640	\$ 887	\$ 1.317	\$ 566	\$ 647	\$ 382	\$ 588	\$ 499	\$ 754
Household	\$ 5.354	\$ 5.787	\$ 4.112	\$ 5.320	\$ 5.410	\$ 5.416	\$ 4.446	\$ 6.812	\$ 3.058	\$ 4.334	\$ 3.974
Kid's Korner	\$ 201	\$ 247	\$ 398	\$ 422	\$ 485	\$ 441	\$ 186	\$ 380	\$ 409	\$ 221	\$ 323
Travel	\$ 624	\$ 608	\$ 505	\$ 559	\$ 1.096	\$ 611	\$ 464	\$ 316	\$ 573	\$ 257	\$ 198

61

OLAP: operatori



drill-across

62

OLAP: operatori

Category	Metrics	Dollar Sales							
		Quarter	Q1 1997	Q2 1997	Q3 1997	Q4 1997	Q1 1998	Q2 1998	Q3 1998
Electronics	\$ 4.383	\$ 817	\$ 827	\$ 4.589	\$ 13.770	\$ 2.977	\$ 4.226	\$ 8.326	
Food	\$ 1.546	\$ 1.310	\$ 1.268	\$ 1.176	\$ 2.676	\$ 1.120	\$ 953	\$ 889	
Gifts	\$ 3.398	\$ 3.893	\$ 4.682	\$ 4.342	\$ 7.879	\$ 4.145	\$ 4.378	\$ 3.645	
Health & Beauty	\$ 1.826	\$ 878	\$ 1.904	\$ 1.434	\$ 2.156	\$ 898	\$ 1.207	\$ 1.404	
Household	\$ 9.314	\$ 8.124	\$ 9.331	\$ 11.614	\$ 17.453	\$ 7.604	\$ 12.898	\$ 12.436	
Kid's Korner	\$ 685	\$ 531	\$ 811	\$ 532	\$ 1.084	\$ 491	\$ 532	\$ 836	
Travel	\$ 603	\$ 1.293	\$ 1.456	\$ 1.145	\$ 1.507	\$ 719	\$ 840	\$ 1.726	

drill-across



Category	Metrics	Quarter		Q1 1997		Q2 1997		Q3 1997		Q4 1997		Q1 1998		Q2 1998		Q3 1998		Q4 1998	
		Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales	Discount	Dollar Sales
Electronics	\$ 0	\$ 4.383	\$ 0	\$ 817	\$ 0	\$ 827	\$ 300	\$ 4.589	\$ 15	\$ 13.770	\$ 0	\$ 2.977							
Food	\$ 25	\$ 1.546	\$ 0	\$ 1.310	\$ 0	\$ 1.268	\$ 38	\$ 1.176	\$ 0	\$ 2.676	\$ 0	\$ 1.120							
Gifts	\$ 31	\$ 3.398	\$ 0	\$ 3.893	\$ 5	\$ 4.682	\$ 0	\$ 4.342	\$ 15	\$ 7.879	\$ 0	\$ 4.145							
Health & Beauty	\$ 0	\$ 1.826	\$ 0	\$ 878	\$ 0	\$ 1.904	\$ 0	\$ 1.434	\$ 229	\$ 2.156	\$ 0	\$ 898							
Household	\$ 0	\$ 9.314	\$ 228	\$ 8.124	\$ 175	\$ 9.331	\$ 35	\$ 11.614	\$ 5	\$ 17.453	\$ 211	\$ 7.604							
Kid's Korner	\$ 0	\$ 685	\$ 0	\$ 531	\$ 32	\$ 811	\$ 40	\$ 532	\$ 0	\$ 1.084	\$ 0	\$ 491							
Travel	\$ 0	\$ 603	\$ 0	\$ 1.293	\$ 200	\$ 1.456	\$ 0	\$ 1.145	\$ 0	\$ 1.507	\$ 0	\$ 719							

63

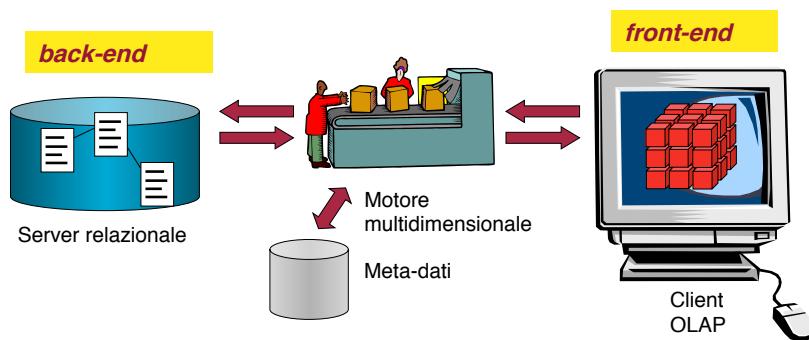
Reportistica semi-statica

- In molti contesti applicativi, è utile un approccio intermedio tra reportistica statica e OLAP: la *reportistica semi-statica*
 - ✓ Un rapporto semi-statico, pur essendo focalizzato su un'insieme di informazioni predefinite, permette all'analista alcuni gradi di libertà, che si concretizzano nella possibilità di eseguire un ristretto insieme di percorsi di navigazione
- Vantaggi:
 - ✓ agli utenti è richiesta una minor competenza sul modello dei dati e sullo strumento di analisi rispetto al caso dell'OLAP
 - ✓ si elimina il rischio di creare risultati d'analisi inconsistenti o scorretti a causa dell'uso improprio dei meccanismi di aggregazione
 - ✓ vincolando i tipi di analisi permessi si evita che l'utente possa involontariamente rallentare il sistema formulando interrogazioni eccessivamente pesanti

64

ROLAP (Relational OLAP)

- Giustificato dall'enorme lavoro svolto in letteratura sul modello relazionale, dalla diffusa esperienza aziendale sull'utilizzo e l'amministrazione di basi di dati relazionali e dall'elevato livello di prestazioni e flessibilità raggiunto dai DBMS relazionali
 - ✓ I dati sono memorizzati su un DBMS relazionale, in forma dettagliata e pre-aggregata
 - ✓ Occorre elaborare tipologie specifiche di schemi che permettano di traslare il modello multidimensionale sul modello relazionale: *schema a stella*
 - ✓ Il problema delle prestazioni porta a *denormalizzazione* per evitare costosi join



65

MOLAP (Multidimensional OLAP)

- Basato su un modello logico ad hoc sul quale i dati e le operazioni multidimensionali possono essere direttamente rappresentati
- I dati vengono fisicamente memorizzati in vettori e l'accesso è di tipo posizionale
 - ✓ Il grosso vantaggio dell'approccio MOLAP rispetto a quello ROLAP è che le operazioni multidimensionali sono realizzabili in modo semplice e naturale, senza necessità di ricorrere a join; le prestazioni risultano pertanto ottime
 - ✓ Non esistendo ancora uno standard per il modello logico multidimensionale, le diverse implementazioni MOLAP hanno veramente poco in comune: in genere, solo l'utilizzo di tecnologie di ottimizzazione specifiche per trattare il problema della sparsità

66

HOLAP (Hybrid OLAP)

- Sistemi di questo tipo combinano in un'unica architettura elementi di ROLAP e MOLAP
 - ✓ Tipicamente i dati di dettaglio sono memorizzati su DBMS relazionale, i pre-aggregati su strutture multidimensionali proprietarie
 - ✓ Oppure, i sottocubi densi sono memorizzati in forma multidimensionale, quelli sparsi in forma relazionale

67

La qualità



La qualità di un processo misura la sua aderenza agli obiettivi degli utenti

- Fattori che caratterizzano la qualità dei dati in un DW:
 1. **Accuratezza:** la conformità tra il valore memorizzato e quello reale.
 2. **Attualità:** il dato memorizzato non è obsoleto.
 3. **Completezza:** non mancano informazioni.
 4. **Consistenza:** la rappresentazione dei dati è uniforme.
 5. **Disponibilità:** i dati sono facilmente disponibili all'utente.
 6. **Tracciabilità:** è possibile risalire alla fonte di ciascun dato.
 7. **Chiarezza:** i dati sono facilmente interpretabili

68



La qualità

- Un ruolo basilare nel raggiungimento degli obiettivi di qualità dei dati è di pertinenza dell'organizzazione aziendale, e potrà essere efficacemente svolto solo mettendo a punto un adeguato e puntuale meccanismo di **certificazione** che individui un ristretto insieme di utenti cui affidare la responsabilità dei dati
- È pertanto un preciso dovere del progettista sensibilizzare i vertici aziendali sull'importanza dell'argomento, e stimolarli affinché mettano a punto un corretto iter di certificazione, opportunamente differenziato per aree aziendali

69

La sicurezza



- La sicurezza dell'informazione è un requisito fondamentale per un sistema, da considerare attentamente nell'ingegneria del software attraverso tutti gli stadi del ciclo di sviluppo
- Il problema della sicurezza è ancora più sentito per i DW, poiché
 - ✓ i DW gestiscono informazione cruciale per il processo decisionale strategico
 - ✓ la multidimensionalità e l'aggregazione introducono addizionali problemi di sicurezza poiché implicitamente consentono inferenze indesiderate sui dati
 - ✓ l'elevata mole di comunicazione che ha luogo nei DW durante l'alimentazione crea specifici problemi relativi alla sicurezza di rete

70

La sicurezza



- Controllo delle autorizzazioni
 - ✓ si svolge all'interno degli strumenti di front-end oppure utilizza i servizi messi a disposizione dai sistemi operativi
- Auditing
 - ✓ le tecniche fornite dai DBMS server non sono in genere sufficienti, e occorre appoggiarsi ai meccanismi implementati dai motori OLAP
- Accesso ai dati per profilo utente
 - ✓ i requisiti di base sono relativi alla mascheratura di interi cubi, di certe loro “fette”, di certe loro misure
 - ✓ in molti casi è necessario nascondere i dati di un cubo oltre un livello di dettaglio assegnato