

Il caso StraSport

(Tratto da: M. Golfarelli, S. Rizzi. Data Warehouse. Teoria e pratica della progettazione.)

La società StraSport si occupa di vendita all'ingrosso di abbigliamento sportivo. Per gestire le proprie vendite si affida ad agenzie distribuite sul territorio italiano, ciascuna delle quali gestisce un certo numero di agenti cui è assegnata una propria zona di lavoro. La maggior parte dei clienti sono grandi catene di negozi con più sedi dislocate sul territorio.

Una vendita è rappresentata da un documento di vendita composto da una o più righe di dettaglio che rappresentano le unità di merce venduta. Gli articoli trattati da StraSport hanno un forte carattere stagionale, il che porta ad un andamento disomogeneo nelle vendite dei diversi prodotti durante l'anno e quindi a un'accentuata fluttuazione del fatturato. L'anno solare comprende principalmente due campagne, Autunno/Inverno e Primavera/Estate, che coprono mesi diversi con riferimento alla raccolta degli ordini e alla fatturazione.

Poco prima dell'inizio di una campagna viene stilato un listino comprendente tutti gli articoli che verranno presentati ai clienti per raccogliere gli ordini e successivamente fatturarli. I prodotti venduti vengono divisi in due categorie, nuovi e classici. I primi sono stati inseriti nella collezione per rinnovare il campionario in funzione delle nuove tendenze dettate dalla moda, i secondi vengono venduti con regolarità per diverse stagioni.

Il budget di esercizio viene stilato una volta all'anno, in estate, per pianificare gli articoli da vendere e gli importi da fatturare in ogni stagione; ovviamente può essere variato durante il corso dell'esercizio. Il budget viene poi affidato alla direzione commerciale e a quella del marketing, che successivamente si occupano di dividere i dati previsionali per cliente e per prodotto.

Pianificazione del data warehouse

Emergono due aree di interesse:

1. **area commerciale:** attenzione rivolta principalmente alle agenzie di vendita
2. **area marketing:** attenzione focalizzata sugli articoli venduti

Queste aree daranno origine a due data mart.

Data mart commerciale

Per il data mart commerciale i fatti di interesse riguardano:

1. **ordinato**
2. **fatturato**
3. **budget**

Ci concentreremo sul fatto ORDINATO.

Analisi e Riconciliazione delle fonti dati

Il sistema informativo di StraSport è costituito da due DB relazionali:

1. **DB Amministrazione:** gestione ordini e fatture
2. **DB Marketing:** programmazione e gestione delle campagne di vendita (budgeting) e dei risultati attesi per agenti di vendita e articoli

Il DB Marketing è il più recente. Sono entrambi realizzati su piattaforma AS/400 ma i sistemi sono scarsamente integrati e presentano sovrapposizioni e conflitti.

Elenco Relazioni del DB Amministrazione

Nome Relazione	Descrizione
CLIENTE	dati anagrafici di base sui clienti
SCHEDA_CLIENTE	informazioni aggiuntive sui clienti
FILIALE	informazioni sulle filiali di ciascun cliente
ARTICOLO	informazioni di base sugli articoli vendibili
SCHEDA_ARTICOLO	informazioni aggiuntive sugli articoli
VENDITORE	agenti di vendita e loro agenzie
MOVIMENTO	testate delle fatture
RIGA_MOVIMENTO	singole righe delle fatture
ORDINE	testate dei documenti di vendita
RIGA_ORDINE	singole righe degli ordini (una tupla per ciascun articolo)

Relazioni

CLIENTE	
<u>CodDitta,CodCliente</u>	PK
RagioneSociale	
Indirizzo	
CAP	
Provincia	
Regione	
MacroArea	
CodiceFiscale	
PartitaIVA	
PrefissoTel	
NumeroTel	
PrefissoTel	
NumeroTel	
CodVenditore	FK: VENDITORE

SCHEDA_CLIENTE	
<u>CodDitta,CodCliente</u>	PK, FK: CLIENTE
DataInizioRapporto	
DataFineRapporto	
ScontoCliente	
ScontoFineAnno	
Solvibilità	
BloccoCommerciale	
BloccoAmministrativo	
GruppoCliente	
ClasseCliente	
ClasseFatturato	
PotenzialitàFatturato	

FILIALE	
<u>CodDitta,CodCliente,CodFiliale</u>	PK, FK: CodDitta,CodCliente REFs CLIENTE
RagioneSociale	
Indirizzo	
Località	
Provincia	
CanaleDistributivo	

ARTICOLO	
<u>CodArticolo</u>	PK
Descrizione	
Colore	
Modello	

SCHEDA_ARTICOLO	
<u>CodArticolo</u>	PK, FK: ARTICOLO
CodiceIVA	
UbicazioneMagazzino	
Fornitore	
Linea	
CostoStandardAcquisto	

VENDITORE	
<u>CodVenditore</u>	PK
Descrizione	
Provvigione	
Agenzia	
Esclusivista	

MOVIMENTO	
<u>Anno,Numero</u>	PK
Data	
CodDitta,CodCliente,CodFiliale	FK: FILIALE

RIGA_MOVIMENTO	
<u>Anno,Numero,NumeroRiga</u>	PK, FK: Anno,Numero REFs MOVIMENTO
CodArticolo	FK: ARTICOLO
Quantità	
ScontoIncondizionato	
Prezzo	

ORDINE	
Anno,Numero,Classe	PK
CodDitta,CodCliente,CodFiliale	FK: FILIALE
CodVenditore	FK: VENDITORE
Pagamento	
DataConsegna	
ScontoSuPagamento	
ScontoIncondizionato	
CausaleAnnullamento	
DataDecorrenzaPagamento	
DataImmissioneOrdine	
QuantitàOrdineTotale	
TipoPagamento	
Stagione	
Classe	

RIGA_ORDINE	
Anno,Numero,Riga	PK, FK: Anno,Numero REFs ORDINE
ClasseOrdine	
CodArticolo	FK: ARTICOLO
QuantitàOrdinata	
DataConsegnaMerce	
CausaleAnnullamento	
CodVenditore	FK: VENDITORE

Analisi delle inconsistenze

In ORDINE:

- il Venditore è ridondante (ogni cliente è seguito da un solo venditore)
- la chiave è una superchiave (la numerazione è indipendente dalla classe d'ordine). Si è trovata l'inconsistenza perché RIGA_ORDINE importa solo i primi due campi della chiave di ORDINE. Verificato anche tramite apposita interrogazione. Classe diventa semplice attributo.

In RIGA_ORDINE:

- il Venditore è ridondante per lo stesso motivo di ORDINE.
- la ClasseOrdine è ridondante perché derivabile dall'ordine cui la riga appartiene.

Si rinomina MOVIMENTO con FATTURA.

NB: Nota che in ORDINE, Anno, Stagione e Classe individuano una CLASSE_ORDINE che individua campagne di particolari classi di ordini. Diventerà un'entità nello schema E/R per potersi sovrapporre al concetto di CAMPAGNA nell'area Marketing (Anno, Stagione).

Ricostruzione dello schema E/R per Reverse Engineering

(schema presentato in aula)

Elenco Relazioni del DB Marketing

Nome Relazione	Descrizione
ARTICOLO	Informazioni di base sugli articoli vendibili
SCHEDA_ARTICOLO	Informazioni aggiuntive sugli articoli
LISTINO	Prezzi di vendita degli articoli
SOTTOSEGMENTO	Sottosegmenti degli articoli (e.g. 'costumi da gara')
SOTTOCATEGORIA	Sottocategorie degli articoli (e.g. 'tessile')
AGENTE	Agenti e loro agenzie
BUDGET_COMMERCIALE	Dati di interesse per il budget commerciale
BUDGET_MARKETING	Dati di interesse per il budget marketing

Relazioni

ARTICOLO	
<u>CodArticolo</u>	PK
Descrizione	
Colore	
Modello	
Tema	
CodSottoSegmento	FK. SOTTOSEGMENTO

SCHEDA_ARTICOLO	
<u>CodArticolo</u>	PK, FK: ARTICOLO
AnnoDiProduzione	
StagioneDiProduzione	
Anno,Stagione,ClasseAttuale	FK: LISTINO
Designer	
PrezzoMedio	
Fornitore	
Target	
Linea	

LISTINO	
<u>Anno,Stagione,Classe,CodArticolo</u>	PK, FK: CodArticolo REFs ARTICOLO
PrezzoUnitario	
ClasseOrdine2	
ClasseOrdine3	

SOTTOSEGMENTO	
<u>CodSottoSegmento</u>	PK
DescrSottoSegmento	
CodSottoCategoria	FK: SOTTOCATEGORIA

SOTTOCATEGORIA	
<u>CodSottoCategoria</u>	PK
DescrSottoCategoria	

AGENTE	
<u>CodAgente</u>	PK
Descrizione	
Provvigione	
Agenzia	
Tipo	

BUDGET_COMMERCIALE	
<u>Anno,Stagione,CodSottoCategoria,CodAgente</u>	PK, FK: CodAgente REFs AGENTE, CodSottoCategoria REFs SOTTOCATEGORIA
PrevisioneResi	
PrevisioneQtaTotale	
PrevisioneAnnullamenti	
PrevisioneSconti	
PrevisioneValoreLordo	

BUDGET_MARKETING	
<u>Anno,Stagione,CodArticolo</u>	PK, FK: CodArticolo REFs ARTICOLO
<u>CodSottoSegmento</u>	FK: SOTTOSEGMENTO
Target	
PrevisioneResi	
PrevisioneAnnullamenti	
PrevisioneScontiFineAnno	
PrevisioneValoreLordo	
PrevisioneQtàTotale	
PrevisioneScontoFattura	

Analisi delle inconsistenze

In SCHEDA_ARTICOLO:

- le informazioni relative al LISTINO non sono utili (LISTINO mantiene già lo storico dei listini in cui un articolo è stato inserito).

In LISTINO:

- denormalizzazione per consentire l'appartenenza di un articolo a più listini in una stessa campagna. Si eliminano perché il numero di listini varia di campagna in campagna e può essere superiore a tre.

In ARTICOLO:

- da un colloquio con gli utenti emerge che gli articolo appartengono a sottosegmenti, a loro volta raggruppati in sottocategorie. Si inseriscono i riferimenti a SOTTOSEGMENTO e a SOTTOCATEGORIA in SOTTOSEGMENTO.

In BUDGET_MARKETING:

- denormalizzazione per dipendenze funzionali CodArticolo → Target e CodArticolo → CodSottoSegmento. Si eliminano.

NB: LISTINO è la relazione che rappresenta l'associazione tra le entità LISTINO e ARTICOLO. L'entità LISTINO nello schema E/R corrisponde alle sole testate di listino (Anno, Stagione, Classe).

Si introduce l'entità CAMPAGNA (Anno, Stagione) riferita nei contesti BUDGET_COMMERCIALE, BUDGET_MARKETING e CLASSE_ORDINE. Quest'ultima viene introdotta per individuare una Classe d'ordine (Anno, Stagione, Classe) coinvolta nella definizione dei listini.

Ricostruzione dello schema E/R per Reverse Engineering
(schema presentato in aula)

Integrazione degli schemi E/R ottenuti per Reverse Engineering

Su indicazione dei responsabili del sistema informativo, in caso di conflitto si privilegerà la soluzione del DB Marketing perché realizzato più recentemente.

Sinonimia:

- VENDITORE e AGENTE risultano sinonimi. Si risolve in AGENTE. La sovrapposizione degli schemi richiede l'unione degli attributi.

Omonimia:

- Fornitore in entrambe le relazioni SCHEDA_ARTICOLO nel DB Amministrazione e nel DB Marketing sono omonimi:
 - o DB Amministrazione: nome del fornitore del prodotto
 - o DB Marketing: nome del produttore

Si rinomina in Produttore nel DB Marketing. La sovrapposizione degli schemi comporta l'unione degli attributi.

Punti di vista differenti:

- Il concetto di CAMPAGNA del DB Marketing è nascosto nel concetto CLASSE_ORDINE nel DB Amministrazione. Si decide di separare gli attributi di CAMPAGNA da CLASSE_ORDINE per consentire la sovrapposizione.
- Le entità AGENTE, ARTICOLO e SCHEDA_ARTICOLO, presenti in entrambi gli schemi rappresentano rispettivamente il medesimo concetto espresso da punti di vista diversi, quindi contengono insiemi di attributi differenti. Sarà necessaria l'unione degli attributi per consentire la sovrapposizione.

Si aggiungono due entità CATEGORIA e SEGMENTO perché rappresentano un livello di aggregazione utile alla classificazione dei dati, quindi al processo decisionale. Questo richiederà un caricamento manuale dei dati.

Con l'introduzione di CATEGORIA, si modella più fedelmente il fatto che il BUDGET_COMMERCIALE viene attualmente redatto per categoria e non sottocategoria (infatti, veniva memorizzata impropriamente un valore di categoria a tale scopo).

Schema E/R riconciliato

(schema presentato in aula)

Progettazione logica del livello riconciliato

CLIENTE	
<u>CodDitta,CodCliente</u>	PK
RagioneSociale	
Indirizzo	
CAP	
Provincia	
Regione	
MacroArea	
CodiceFiscale	
PartitaIVA	
PrefissoTel	
NumeroTel	
PrefissoTel	
NumeroTel	
CodAgente	FK: AGENTE
DataInizioRapporto	
DataFineRapporto	
ScontoCliente	
ScontoFineAnno	
Solvibilità	
BloccoCommerciale	
BloccoAmministrativo	
GruppoCliente	
ClasseCliente	
ClasseFatturato	
PotenzialitàFatturato	

AGENTE	
<u>CodAgente</u>	PK
Descrizione	
Provvigione	
Agenzia	
Esclusivista	
Tipo	

FILIALE	
<u>CodDitta,CodCliente,CodFiliale</u>	PK, FK: CodDitta,CodCliente REFs CLIENTE
RagioneSociale	
Indirizzo	
Località	
Provincia	
CanaleDistributivo	

FATTURA	
<u>Anno,Numero</u>	PK
Data	
CodDitta,CodCliente,CodFiliale	FK: FILIALE

RIGA_FATTURA	
<u>Anno,Numero,NumeroRiga</u>	PK, FK: Anno,Numero REFs FATTURA
CodArticolo	FK: ARTICOLO
Quantità	
ScontoIncondizionato	
Prezzo	

ORDINE	
<u>Anno,Numero</u>	PK
Stagione,Classe	FK: Anno,Stagione,Classe REFs CLASSE_ORDINE
CodDitta,CodCliente,CodFiliale	FK: FILIALE
Pagamento	
DataConsegna	
ScontoSuPagamento	
ScontoIncondizionato	
CausalePagamento	
DataDecorrenzaPagamento	
DataEmissioneOrdine	
QuantitàOrdineTotale	
TipoPagamento	

RIGA_ORDINE	
<u>Anno,Numero,NumeroRiga</u>	PK, FK: Anno,Numero REFs ORDINE
CodArticolo	FK: ARTICOLO
QuantitàOrdinata	
DataConsegnaMerce	
CausaleAnnullamento	

CLASSE_ORDINE	
<u>Anno,Stagione,Classe</u>	PK, FK: Anno,Stagione REFs CAMPAGNA

CAMPAGNA	
<u>Anno,Stagione</u>	PK

ARTICOLO	
<u>CodArticolo</u>	PK
Descrizione	
Colore	
Modello	
Tema	
CodiceIVA	
AnnoProduzione	
StagioneProduzione	
UbicazioneMagazzino	
Designer	
PrezzoMedio	
Produttore	
Fornitore	
Target	
Linea	
CostoStandardAcquisto	
CodSottoSegmento	FK: SOTTOSEGMENTO

LISTINO	
<u>Anno,Stagione,Classe,CodArticolo</u>	PK, FK: CodArticolo REFs ARTICOLO FK: Anno,Stagione,Classe REFs CLASSE_ORDINE
Prezzo	

CATEGORIA	
<u>CodCategoria</u>	PK
Descrizione	

SOTTOCATEGORIA	
<u>CodSottoCategoria</u>	PK
Descrizione	
CodCategoria	FK: CATEGORIA

SEGMENTO	
<u>CodSegmento</u>	PK
Descrizione	
CodSottoCategoria	FK: SOTTOCATEGORIA

SOTTOSEGMENTO	
<u>CodSottoSegmento</u>	PK
Descrizione	
CodSegmento	FK: SEGMENTO

BUDGET_COMMERCIALE	
<u>Anno,Stagione,CodAgente,CodCategoria</u>	PK, FK: CodAgente REFs AGENTE FK: Anno,Stagione REFs CAMPAGNA FK: CodCategoria REFs CATEGORIA
PrevResi	
PrevAnnullamenti	
PrevSconti	
PrevValoreLordo	
PrevQtàTotale	

BUDGET_MARKETING	
<u>Anno,Stagione,CodArticolo</u>	PK, FK: CodArticolo REFs ARTICOLO FK: Anno,Stagione REFs CAMPAGNA
PrevResi	
PrevAnnullamenti	
PrevScontiFineAnno	
PrevValoreLordo	
PrevQtàTotale	
PrevScontoFattura	

Mapping tra le relazioni degli schemi sorgenti e le relazioni dello schema riconciliato

Esempi:

// Mapping dei clienti dal DB Commerciale

```
CREATE VIEW CLIENTE AS
    SELECT CL.CLIDIT,CL.CLICOD, CL.RAGSOC,..., SC.PREFTEL, SC.TEL,...
    FROM CLIENTE AS CL, SCHEDA_CLIENTE AS SC
    WHERE CL.CLIDIT=SC.CLIDIT
        AND CL.CLICOD=SC.CLICOD
```

// Mapping delle campagne dal DB Marketing

```
CREATE VIEW CAMPAGNE AS
    SELECT ANNO,STAGIONE
    FROM BUDGET_COMMERCIALE
    UNION
    SELECT ANNO,STAGIONE
    FROM BUDGET_MARKETING
```

Analisi dei requisiti utente

Glossario:

Fatto	Possibili dimensioni	Possibili misure	Storicità
ORDINATO	agente, cliente, sottocategoria dell'articolo, data ordine, data pagamento, data consegna merce	qtà ordinata, prezzo, sconti, resi, ordinato netto	5 anni
FATTURATO	agente, cliente, sottocategoria dell'articolo, data fattura	qtà fatturata, prezzo, sconti	5 anni
BUDGET	anno, stagione, agente, categoria dell'articolo	dati previsionali	5 anni

Carico di lavoro preliminare:

Fatto	Interrogazione
ORDINATO	Quantità totale ordinata annualmente per una certa categoria di articoli. Differenza media tra data consegna merce e data del pagamento per cliente. Ordinato netto totale per agente, cliente e sottocategoria di articolo.
FATTURATO	Quantità totali fatturate per ciascuna sottocategoria di articoli nell'ultimo anno. Per un cliente, incassi relativi alle diverse categorie di articoli per un dato mese. Riepilogo annuale degli incassi per regione del cliente
BUDGET	Riepilogo delle previsioni delle quantità totali vendute per anno e categoria articolo. Previsioni annullamenti e resi per l'anno in corso per agente e stagione.

Interventi

Introduzione dell'attributo RESI per ciascuna riga d'ordine.

Si concorda un intervallo di aggiornamento pari a una settimana per ORDINATO e FATTURATO e un intervallo annuale per il BUDGET.