

Università della Calabria
Corso di Laurea Magistrale in
Data Science per le Strategie Aziendali

Anno Accademico 2023-2024

Sviluppo del Data Warehouse
di un'azienda alimentare

Sommario

Sommario

Sommario	2
1. Introduzione	3
2. Il dominio applicativo	4
3. Pianificazione del data warehouse	4
3.1 Il data mart commerciale	5
3.1.1 Analisi e riconciliazione delle fonti dati	5
3.1.2 Analisi dei requisiti utente	15
3.1.5 Progettazione logica	28
3.1.6 Progettazione dell'alimentazione: dal riconciliato al derivato	31
3.1.7 Cubi OLAP	35
3.1.8 Reportistica	42
3.1.9 Dashboard	46

1. Introduzione

Il *data warehousing* è un sistema di supporto alle decisioni, il cui scopo è quello di permettere al decisore umano di compiere scelte più consapevoli fornendo una visione chiara e profonda dei dati disponibili. Il *data warehousing* nasce in risposta all'esigenza di estrarre informazione da una quantità elevata di dati con un sistema dedicato esclusivamente all'analisi e tale da non interferire con le attività transazionali di routine.

Il *data warehouse* (DW), protagonista del *data warehousing*, è definito come una collezione di dati di supporto per il processo decisionale. È necessario che il DW sia costruito intorno ai *concetti* di interesse aziendali (clienti, vendite, prodotti, ad esempio), e non intorno alle diverse aree del dominio aziendale. Inoltre, il DW deve fornire una visione unica, coerente e consistente di tutte le sorgenti di dati disponibili, che possono consistere in DB di proprietà dell'azienda oppure in collezioni di dati estratte da sistemi informativi esterni. Infine, il DW deve essere aggiornato continuamente e deve conservare i dati per lunghi periodi, in modo da permettere analisi riguardanti l'evoluzione temporale relativa anche a diversi anni.

Il DW è un contenitore centrale e globale dei dati aziendali, che viene poi utilizzato come sorgente di dati per *più piccoli* DW locali, relativi a un determinato settore dell'azienda, detti *data mart*. In questo progetto, ci limiteremo alla progettazione di un solo *data mart* indipendente.

2. Il dominio applicativo

La società di riferimento opera nel mercato alimentare e si occupa della produzione e della vendita al dettaglio di prodotti alimentari. Gli articoli vendibili si dividono in: pasta, sottoli, sughi, spezie. L'azienda si rifornisce di materie prime necessarie ai processi produttivi effettuando ordini a diversi fornitori. Il sistema informativo utilizzato dalla società tiene traccia di tali ordini e gestisce il processo di vendita.

3. Pianificazione del data warehouse

Nella fase di pianificazione del data warehouse, si è scelto di utilizzare un'architettura a tre livelli, con il sostanziale vantaggio di prevedere un livello dei dati riconciliati che alimentano il data warehouse primario. I dati che popolano il data warehouse primario sono stati sottoposti alle attività degli strumenti ETL, pertanto sono coerenti, consistenti e dettagliati.

I settori aziendali di interesse sono diversi; in ogni caso, si è scelto di iniziare con la progettazione del data mart dell'area commerciale. Avere una visione chiara e profonda dell'universo commerciale in tempi brevi è, infatti, un requisito che la società reputa primario. Inoltre, il data mart dell'area commerciale è quello che riveste il ruolo più strategico e il suo sviluppo, con i conseguenti benefici che ci aspetta di osservare, può tenere alto l'interesse e l'investimento dell'azienda durante l'intero processo.

Per la costruzione di un sistema di DW è possibile seguire, sostanzialmente, due approcci: lasciare che la progettazione venga guidata dai dati, oppure dai requisiti.

Se si sceglie di seguire il primo percorso, la prima fase è quella di analisi e riconciliazione dei dati. Tutte le fasi successive saranno effettuate tenendo bene a mente la natura dei dati disponibili, in modo da produrre specifiche compatibili. Inoltre, le fasi di analisi dei requisiti utente e della progettazione concettuale possono essere svolte in contemporanea.

Seguire, invece, l'approccio guidato dai requisiti prevede di svolgere l'analisi dei requisiti utente all'inizio del processo; questo potrebbe rendere difficoltosa la successiva fase di analisi e riconciliazione delle sorgenti dati, oltre a non rendere necessariamente possibile il soddisfacimento di tutte le esigenze del cliente per le quali, nella fase iniziale, si erano proposte delle ipotetiche soluzioni.

Vista l'elevata qualità e la relativa semplicità dei DB presenti nel sistema informativo dell'azienda, si è scelto di seguire un approccio guidato dai dati. Pertanto, la prima fase è stata quella di analisi e riconciliazione delle fonti dati.

3.1 Il data mart commerciale

3.1.1 Analisi e riconciliazione delle fonti dati

Questa fase ha l'obiettivo di definire e documentare lo schema del livello dei dati riconciliati. I dati contenuti nel DW sono ricavati da sorgenti di dati interne o esterne all'azienda e, potenzialmente, molto diverse tra loro, sia dal punto di vista della tecnologia che da quello del modello utilizzato per rappresentare la realtà aziendale. Le varie sorgenti possono essere correlate, sovrapposte oppure totalmente indipendenti.

In questa fase del processo, è necessario:

- analizzare le sorgenti singolarmente e comprenderne lo schema;
- valutare la presenza di inconsistenze;
- sanare incongruenze tra le sorgenti;
- osservare l'eventuale presenza di concetti comuni o simili da unificare;
- eliminare porzioni ritenute non utili.

È indispensabile che a questa fase prendano parte, oltre ai progettisti, gli amministratori dei DB, già a conoscenza del significato degli schemi. In uscita, si ottiene lo schema dei dati riconciliati.

Le analisi delle fonti dei dati vengono effettuate sulla porzione del sistema informativo aziendale che riguarda l'area commerciale. Il sistema informativo contiene due DB di interesse: il DB Forniture e il DB Vendite.

I due DB sono stati introdotti nel sistema aziendale contemporaneamente e non presentano conflitti di grossa entità. Di seguito sono riportate le relazioni presenti nei due DB e i relativi schemi E-R, relazionale e logico.

Tabella 3.1 Relazioni nel DB Forniture

<i>Relazioni</i>	<i>Descrizione</i>	<i>Attributi</i>
Indirizzo	Indirizzo del fornitore	ID (identificativo, INT) Via (via, VARCHAR(50)) Numero (numero civico, INT) Stato (stato, VARCHAR(50)) Regione (regione, VARCHAR(50)) Città (città, VARCHAR(50))
Fornitore	Dati del fornitore	Iva (numero di partita iva, VARCHAR(12)) Nome (nome del fornitore, VARCHAR(50)) Email (email del fornitore VARCHAR(50)) NumeroTelefonico (numero di telefono del fornitore, BIGINT)
OrdineFornitore	Ordine a un fornitore	ID (identificativo del fornitore, INT) IvaFornitore (VARCHAR(12))
MateriaPrima	Informazioni sulle materie prime acquistate dall'azienda	ID (identificativo della materia prima, INT) Nome (nome della materia prima, VARCHAR(50)) Lotto (identificativo del lotto di provenienza, VARCHAR(50)) DataScadenza (data di scadenza della materia prima, DATETIME) Tipologia (categoria a cui la materia prima appartiene, VARCHAR(50))

Prodotto	Informazioni sui prodotti vendibili dall'azienda	ID (identificativo del prodotto, INT) Tipologia (tipologia di prodotto, VARCHAR(50)) Nome (nome del prodotto, VARCHAR(100))
Magazzino	Prodotto presente in magazzino	ID (identificativo del prodotto presente in magazzino, INT) Nome (nome del prodotto, VARCHAR(100)) Lotto (identificativo del lotto a cui il prodotto appartiene, VARCHAR(50)) DataProduzione (data di produzione del prodotto, DATETIME) DataScadenza (data di scadenza del prodotto, DATETIME)
Fornitura	Documento contenente informazioni su un ordine di fornitura con relativi importo e quantità	OrdineFornitura (identificativo dell'ordine della fornitura, INT) MateriaPrimaFornitura (identificativo della materia prima a cui si riferisce la fornitura, INT) Data (data di richiesta della fornitura, DATETIME) PrezzoUnitario (prezzo unitario della materia prima, DOUBLE) Importo (importo totale della fornitura, DOUBLE) Quantità (quantità di materia prima richiesta, INT)
Contenuto	Informazioni sulle materie prime che costituiscono un prodotto, e viceversa	MateriaContenuta (identificativo della materia prima, INT) ProdottoContenuto (identificativo del prodotto, INT)

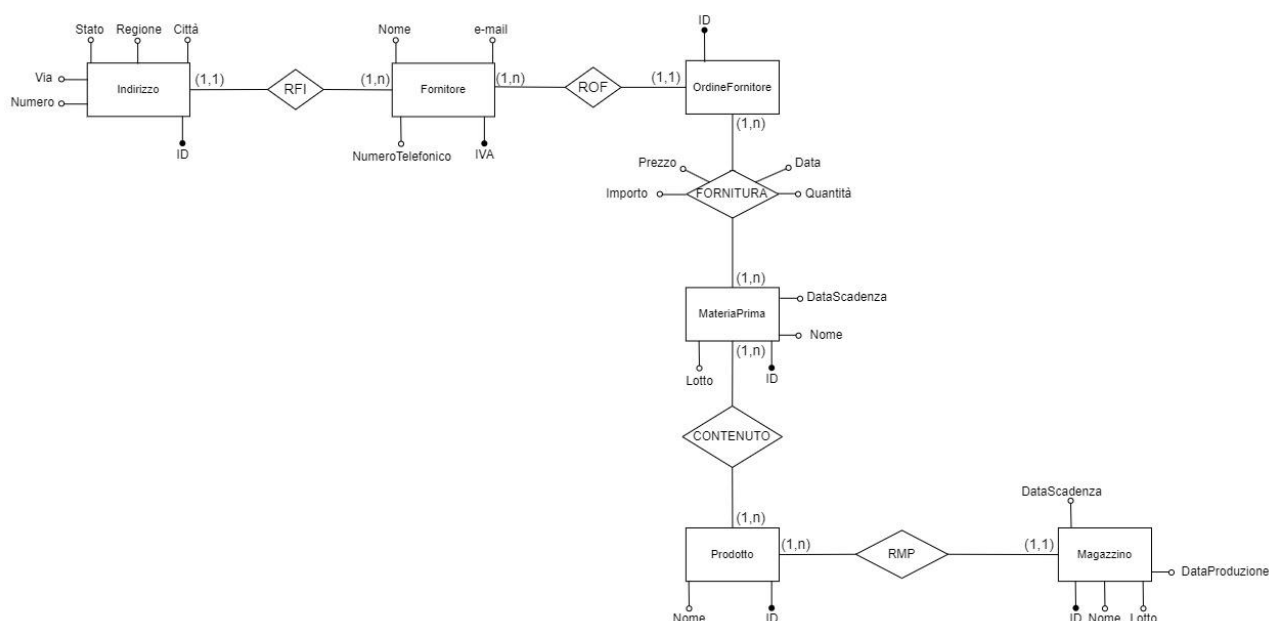


Figura 3.1 Schema Entità – Relazione del DB Forniture

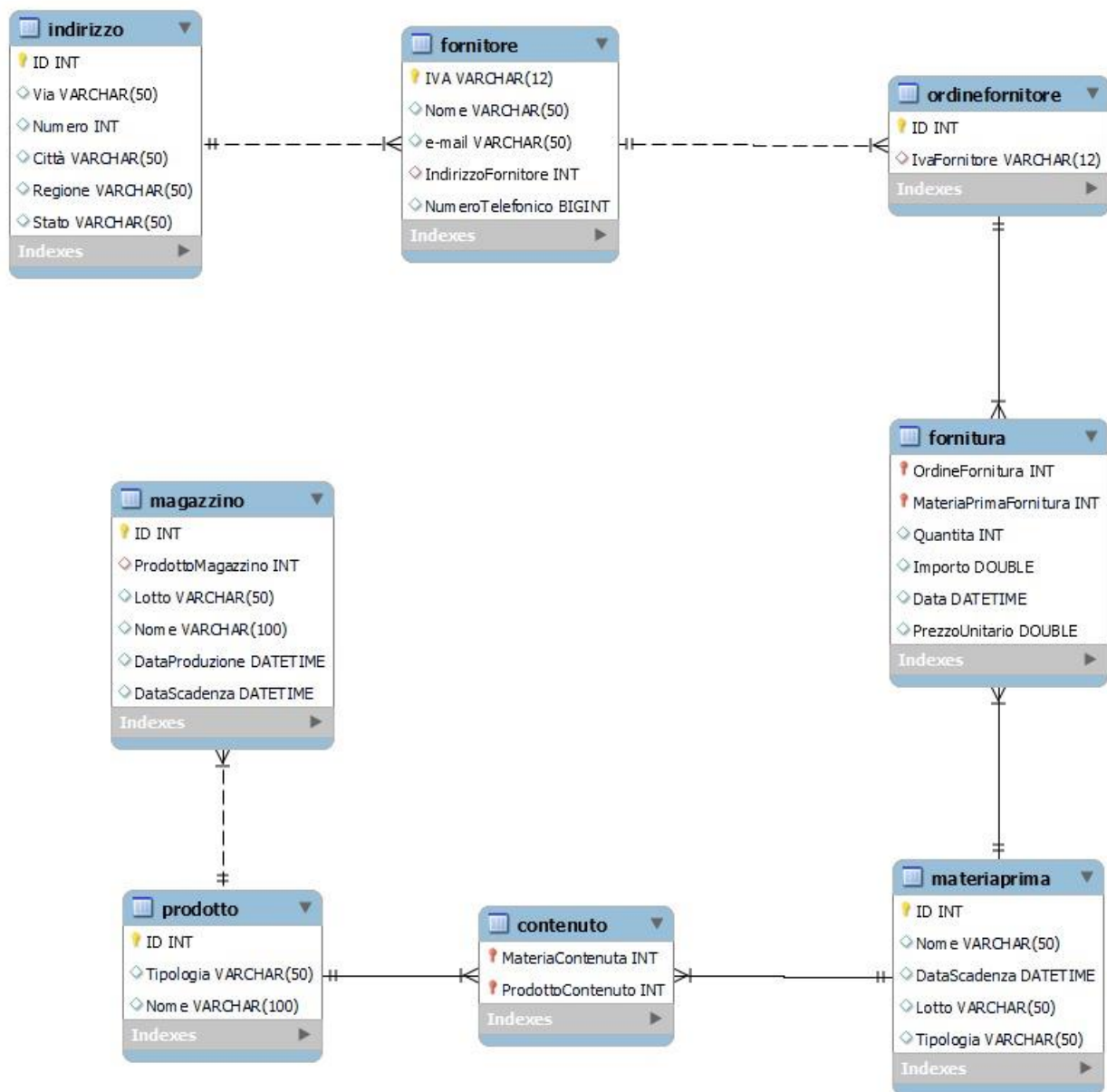


Figura 3.2 Schema relazionale del DB Forniture

Schema logico del DB Forniture

Indirizzo(ID, Via, Numero, Città, Regione, Stato)

Fornitore(IVA, Nome, email, IndirizzoFornitore, NumeroTelefonico)

OrdineFornitore(ID, IvaFornitore: **Fornitore**)

MateriaPrima(ID, Nome, DataScadenza, Lotto, Tipologia)

Prodotto(ID, Tipologia, Nome)

Magazzino(ID, ProdottoMagazzino: **Prodotto**, Lotto, Nome, DataProduzione, DataScadenza)

Fornitura(OrdineFornitura: **OrdineFornitore**, MateriaPrimaFornitura: **MateriaPrima**, Quantità Importo, Data, PrezzoUnitario)

Contenuto(MateriaContenuta: **MateriaPrima**, ProdottoContenuto: **Prodotto**)

Tabella 3.2 Relazioni nel DB Vendite

<i>Relazioni</i>	<i>Descrizione</i>	<i>Attributi</i>
Indirizzo	Indirizzo del cliente	ID (identificativo, INT) Via (via, VARCHAR(50)) Numero (numero civico, INT) Stato (stato, VARCHAR(50)) Regione (regione, VARCHAR(50)) Città (città, VARCHAR(50))
Cliente	Dati del cliente	ID (identificativo del cliente, INT) Nome (nome del cliente, VARCHAR(50)) IndirizzoCliente (identificativo dell'indirizzo del cliente, INT) Email (email del cliente, VARCHAR(50)) NumeroTelefonico (numero di telefono del cliente, BIGINT) Sconto (percentuale di sconto a tutti gli ordini che il cliente effettua, DOUBLE)
OrdineCliente	Ordine di un cliente	ID (identificativo dell'ordine del cliente, INT) DataConsegna (data di consegna dell'ordine, DATETIME) ClienteOrdine (identificativo del cliente che ha effettuato l'ordine, INT)
Prodotto	Informazioni sui prodotti vendibili dall'azienda	ID (identificativo del prodotto, INT) Nome (nome del prodotto, VARCHAR(50)) Tipologia (tipologia di prodotto, VARCHAR(100))
Vendita	Documento contenente le informazioni relative alla vendita di un prodotto, con relativi importo e quantità	OrdineVendita (identificativo dell'ordine del cliente, INT) ProdottoVendita (identificativo del prodotto richiesto, INT) Data (data della vendita, DATETIME) PrezzoUnitario (prezzo unitario del prodotto, DOUBLE) Importo (importo totale della vendita, DOUBLE) Quantità (quantità di prodotto richiesta, INT)

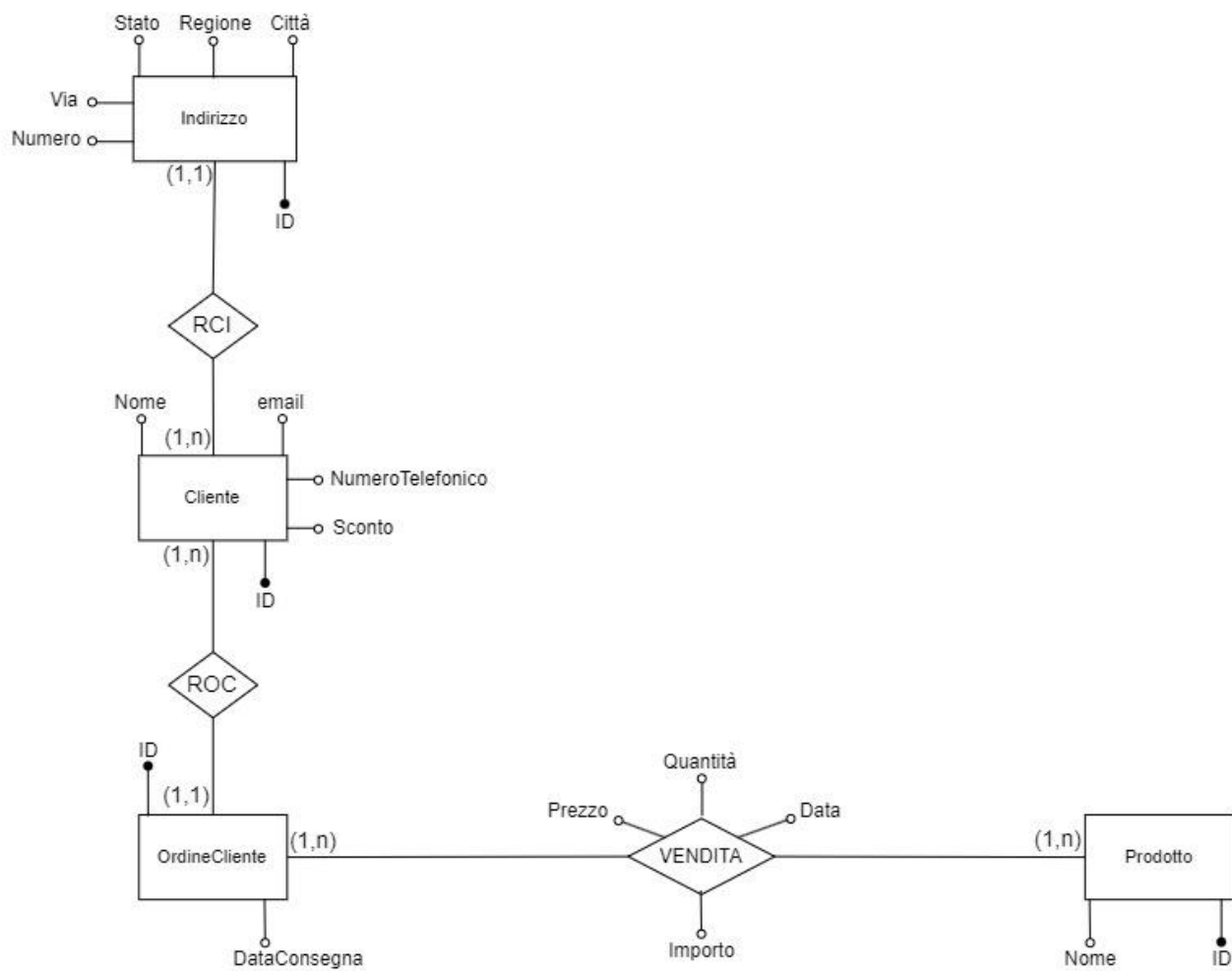


Figura 3.3 Schema Entità – Relazione del DB Vendite

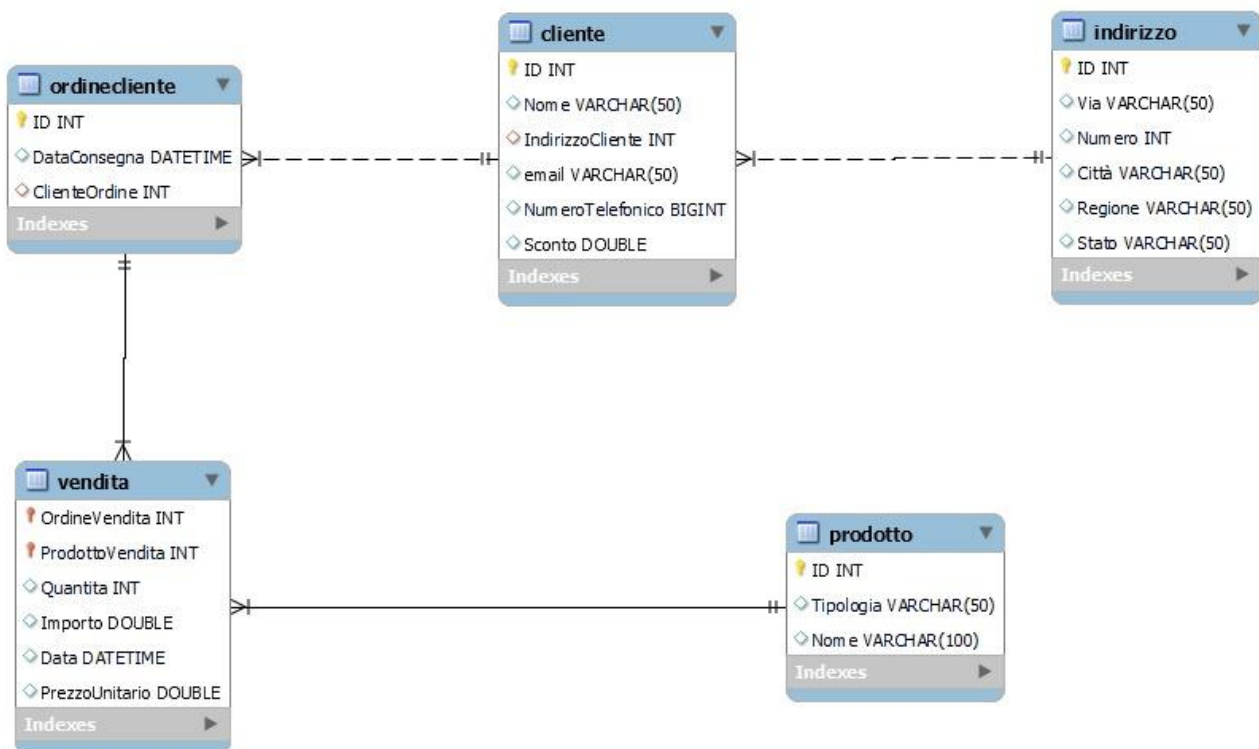


Figura 3.4 Schema relazionale del DB Vendite

Schema logico del DB Vendite

Indirizzo (<u>ID</u> , Via, Numero, Città, Regione, Stato)
Cliente (<u>ID</u> , Nome, email, NumeroTelefonico, Sconto)
OrdineCliente (<u>ID</u> , DataConsegna, ClienteOrdine: Cliente)
Prodotto (<u>ID</u> , Nome, Tipologia)
Vendita (<u>OrdineVendita</u> : OrdineCliente , <u>ProdottoVendita</u> : Prodotto , Quantità, Importo, Data, PrezzoUnitario)

Dalla comparazione dei due schemi emerge la presenza di concetti comuni:

- la relazione **Indirizzo** è comune a entrambi i DB;
- la relazione **Prodotto** è comune a entrambi i DB.

A partire dai precedenti schemi, si procede all'integrazione delle due sorgenti, sovrapponendo i concetti comuni.

Tabella 3.3 Relazioni nel DB riconciliato

Relazioni	Descrizione	Attributi
Indirizzo	Indirizzo (del fornitore o del cliente)	ID (identificativo, INT) Via (via, VARCHAR(50)) Numero (numero civico, INT) Stato (stato, VARCHAR(50)) Regione (regione, VARCHAR(50)) Città (città, VARCHAR(50))
Cliente	Dati del cliente	ID (identificativo del cliente, INT)

		Nome (nome del cliente, VARCHAR(50)) IndirizzoCliente (identificativo dell'indirizzo del cliente, INT) Email (email del cliente, VARCHAR(50)) NumeroTelefenico (numero di telefono del cliente, BIGINT) Sconto (percentuale di sconto a tutti gli ordini che il cliente effettua, DOUBLE)
OrdineCliente	Ordine di un cliente	ID (identificativo dell'ordine del cliente, INT) DataConsegna (data di consegna dell'ordine, DATETIME) ClienteOrdine (identificativo del cliente che ha effettuato l'ordine, INT)
Fornitore	Dati del fornitore	Iva (numero di partita iva, VARCHAR(12)) Nome (nome del fornitore, VARCHAR(50)) Email (email del fornitore VARCHAR(50)) NumeroTelefonico (numero di telefono del fornitore, BIGINT)
OrdineFornitore	Ordine a un fornitore	ID (identificativo del fornitore, INT) IvaFornitore (VARCHAR(12))
MateriaPrima	Informazioni sulle materie prime acquistate dall'azienda	ID (identificativo della materia prima, INT) Nome (nome della materia prima, VARCHAR(50)) Lotto (identificativo del lotto di provenienza, VARCHAR(50))
Prodotto	Informazioni sui prodotti vendibili dall'azienda	ID (identificativo del prodotto, INT) Tipologia (tipologia di prodotto, VARCHAR(50)) Nome (nome del prodotto, VARCHAR(100))
Magazzino	Prodotto presente in magazzino	ID (identificativo del prodotto presente in magazzino, INT) Nome (nome del prodotto, VARCHAR(100)) Lotto (identificativo del lotto a cui il prodotto appartiene, VARCHAR(50)) DataProduzione (data di produzione del prodotto, DATETIME) DataScadenza (data di scadenza del prodotto, DATETIME)
Vendita	Documento contenente le informazioni relative alla vendita di un prodotto, con relativi importo e quantità	OrdineVendita (identificativo dell'ordine del cliente, INT) ProdottoVendita (identificativo del prodotto richiesto, INT) Data (data della vendita, DATETIME) PrezzoUnitario (prezzo unitario del prodotto, DOUBLE) Importo (importo totale della vendita, DOUBLE) Quantità (quantità di prodotto richiesta, INT)
Contenuto	Informazioni sulle materie prime che costituiscono un prodotto, e viceversa	MateriaContenuta (identificativo della materia prima, INT) ProdottoContenuto (identificativo del prodotto, INT)

Fornitura	Documento contenente informazioni su un ordine di fornitura con relativi importo e quantità	<p><u>OrdineFornitura</u> (identificativo dell'ordine della fornitura, INT)</p> <p><u>MateriaPrimaFornitura</u> (identificativo della materia prima a cui si riferisce la fornitura, INT)</p> <p>Data (data di richiesta della fornitura, DATETIME)</p> <p>PrezzoUnitario (prezzo unitario della materia prima, DOUBLE)</p> <p>Importo (importo totale della fornitura, DOUBLE)</p> <p>Quantità (quantità di materia prima richiesta, INT)</p>
-----------	---	---

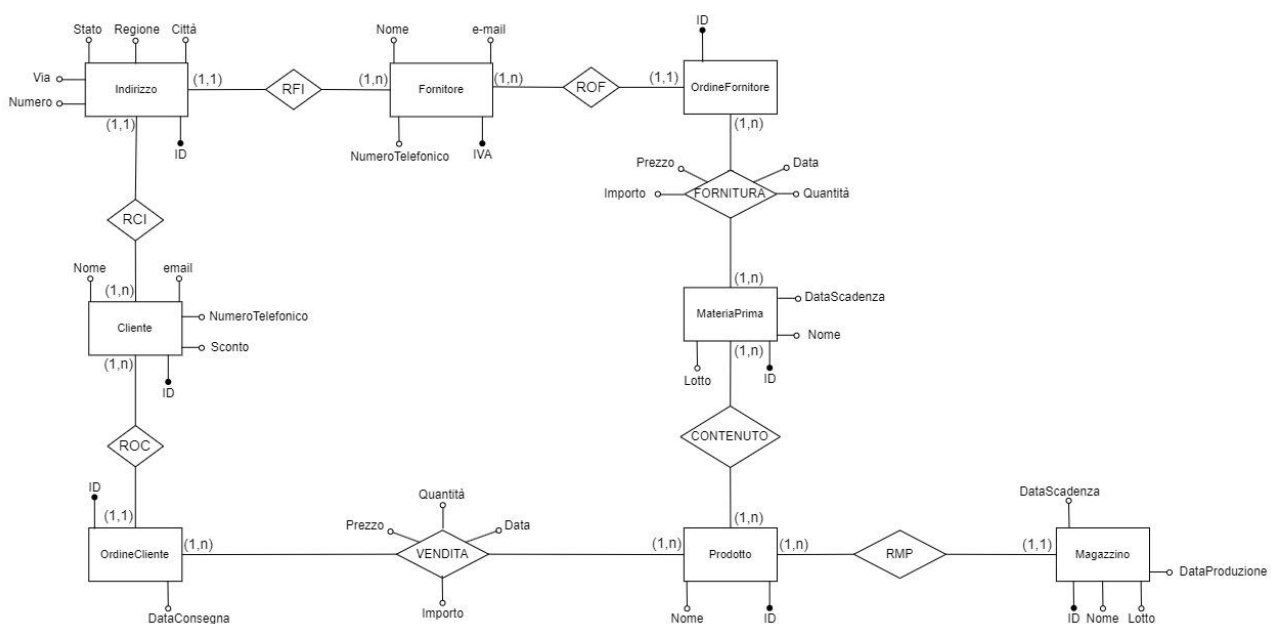


Figura 3.5 Schema Entità – Relazione del DB riconciliato

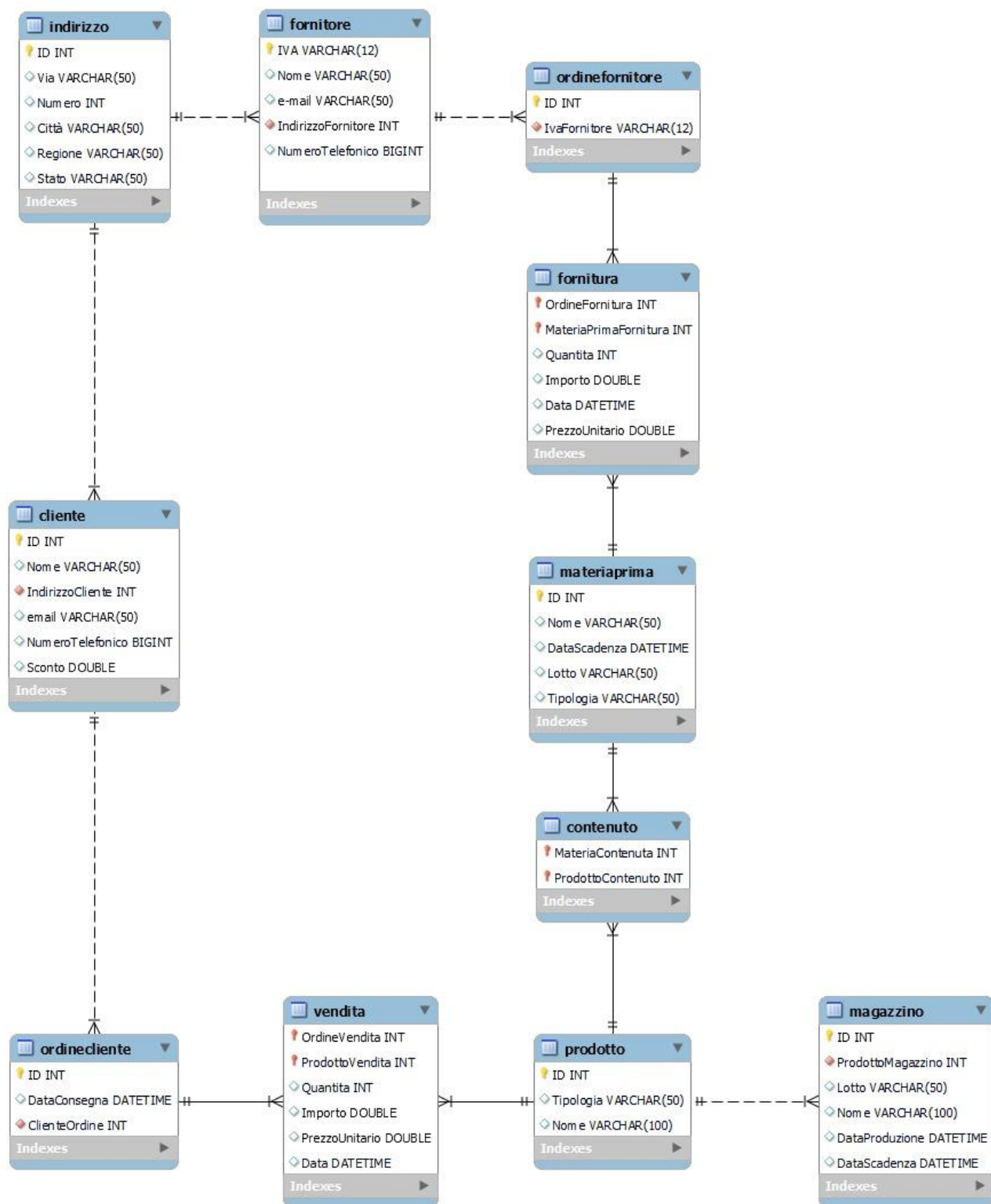


Figura 3.6 Schema relazionale del DB riconciliato

Schema logico del DB riconciliato

Indirizzo(ID, Via, Numero, Città, Regione, Stato)

Cliente(ID, Nome, email, NumeroTelefonico, Sconto)

OrdineCliente(ID, DataConsegna)

Fornitore(IVA, Nome, email, IndirizzoFornitore, NumeroTelefonico)

OrdineFornitore(ID, IvaFornitore: **Fornitore**)

MateriaPrima(ID, Nome, DataScadenza, Lotto, Tipologia)

Prodotto(ID, nome)

Magazzino(ID, ProdottoMagazzino: **Prodotto**, Lotto, Nome, DataProduzione, DataScadenza)

Vendita(OrdineVendita: **OrdineCliente**, ProdottoVendita: **Prodotto**, Quantità, Importo, Data, PrezzoUnitario)

Contenuto(MateriaContenuta: **MateriaPrima**, ProdottoContenuto: **Prodotto**)

Fornitura(OrdineFornitura: **OrdineFornitore**, MateriaPrimaFornitura: **MateriaPrima**, Quantità Importo, Data, PrezzoUnitario)

3.1.2 Analisi dei requisiti utente

La fase di analisi dei requisiti utente è particolarmente delicata, perché influenza tutte le decisioni riguardo alla progettazione concettuale, alle applicazioni da implementare, alla progettazione dell'alimentazione.

La fonte principale di informazione sono le interviste agli utenti finali del data mart. Le interviste devono essere condotte con attenzione: eventuali incomprensioni relative al lessico o al linguaggio tra i progettisti e gli intervistati è spesso causa, alla fine del processo, dell'insoddisfazione del cliente.

In questa fase, il progettista deve raccogliere i requisiti degli utenti finali e tradurli, tenendo a mente gli obiettivi strategici, in specifiche sui fatti da modellare e sul carico di lavoro. In particolare, deve

- individuare i fatti di interesse
- individuare possibili dimensioni di analisi e misure
- scegliere la granularità.

Dalle interviste condotte agli utenti finali del data mart, è emerso l'interesse per due fatti di interesse: fornitura e vendita.

Le richieste degli utenti hanno portato all'aggiunta di ulteriori livelli di dettaglio nella dimensione di analisi temporale.

Si concorda con gli utenti un intervallo di aggiornamento settimanale.

Tabella 3.4 Glossario dei requisiti utente

<i>Fatto</i>	<i>Possibili dimensioni</i>	<i>Possibili misure</i>	<i>Storicità</i>
Fornitura	Materia prima, data, ordine	Quantità, importo, prezzo unitario	2 anni
Vendita	Prodotto, ordine, data	Quantità, importo, prezzo unitario	2 anni

Tabella 3.5 Campione di carico di lavoro preliminare

Fatto	Interrogazioni
Fornitura	<ul style="list-style-type: none">• Importo speso per l'acquisto di materie prime nei diversi mesi per città• Quantità di materie prime acquistate nelle diverse regioni per semestre• Il prezzo unitario massimo di una specifica tipologia di materia prima acquistata in una città per anno• Prezzo unitario medio delle materie prime appartenenti ad una specifica tipologia acquistate in un mese.
Vendita	<ul style="list-style-type: none">• Regione nella quale l'importo ricavato dalla vendita di una tipologia di prodotto in uno specifico mese è risultato inferiore• Quantità di prodotti venduti in Italia per mese• Massimo prezzo unitario di uno specifico prodotto venduto in una città per giorno

3.1.3 Progettazione dell'alimentazione: dalle sorgenti al riconciliato

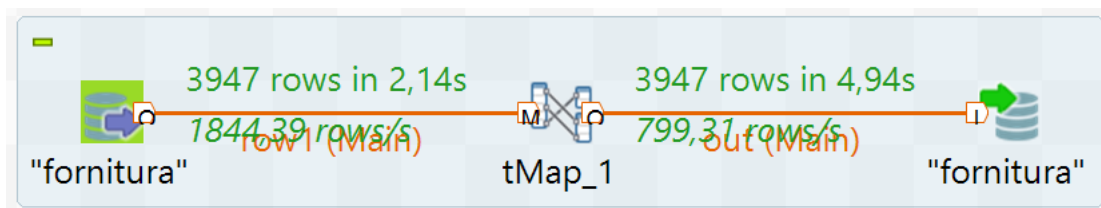
Il processo di alimentazione comprende le seguenti procedure.

- **Estrazione:** i dati vengono estratti dalle sorgenti. Se non è mai stata effettuata una estrazione, tutti i dati vengono estratti. Se, invece, si tratta di un aggiornamento periodico, vengono estratti soltanto i dati che sono stati inseriti o modificati dopo l'ultima estrazione.
- **Pulitura:** viene corretta la presenza di dati duplicati, inconsistenze, dati mancanti, valori impossibili.
- **Trasformazione:** vengono uniformati i formati di memorizzazione e le unità di misura; inoltre, si stabiliscono corrispondenze tra campi equivalenti in sorgenti diverse.
- **Caricamento:** i dati vengono riscritti, integralmente o parzialmente.

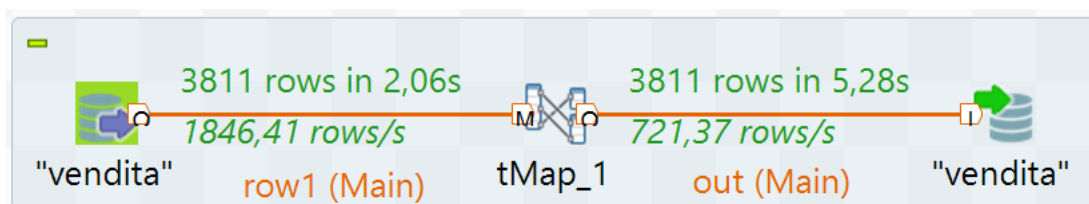
Il data mart che si sta progettando deve essere popolato per la prima volta. Per questo motivo, si effettua una estrazione statica dei dati dalle sorgenti e, in seguito, un'operazione di *refresh*.

Di seguito, i *job* utilizzati per popolare il DB riconciliato, in seguito alla connessione con i DB su MySQL.

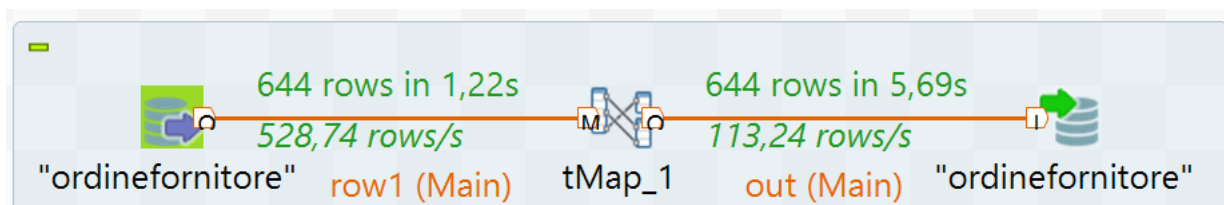
Fornitura



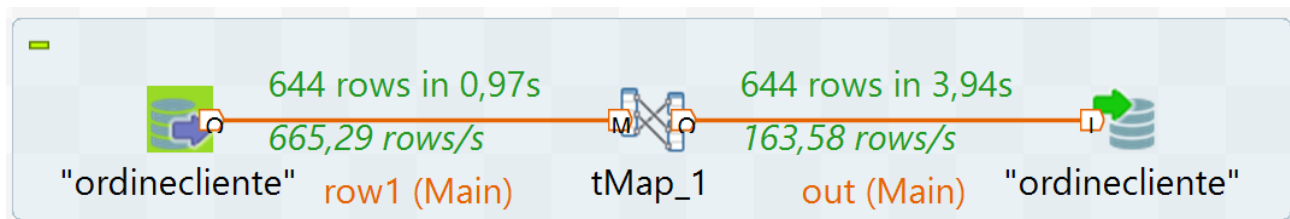
Vendita



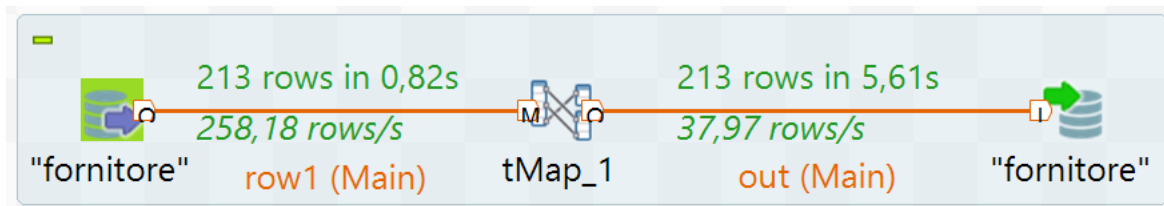
OrdineFornitore



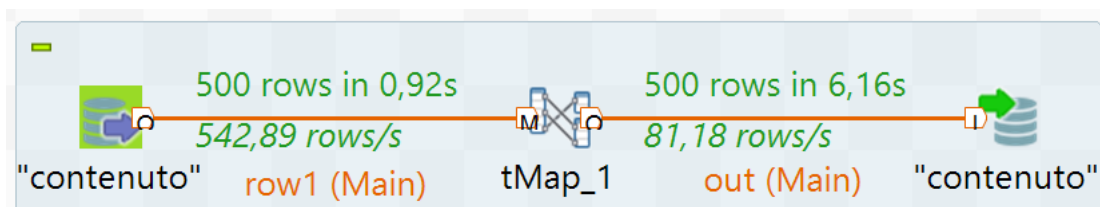
OrdineCliente



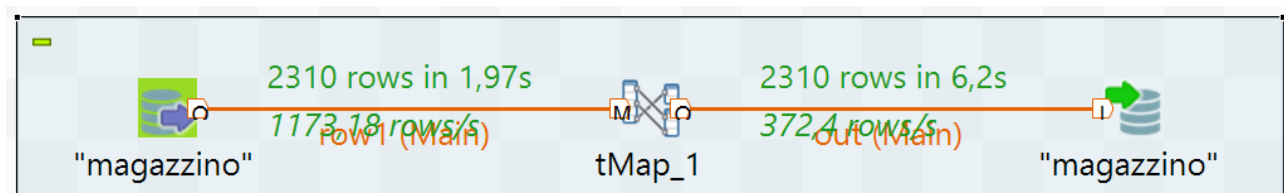
Fornitore



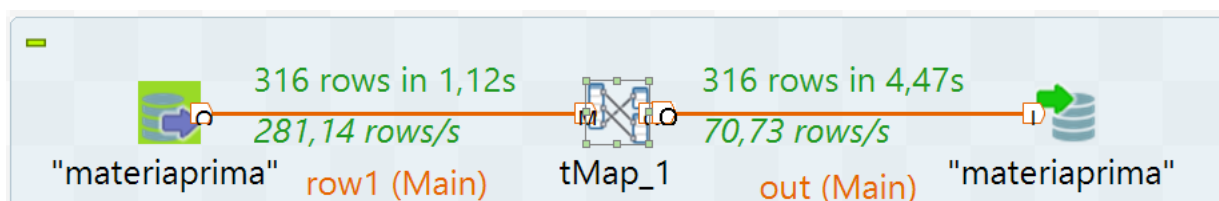
Contenuto



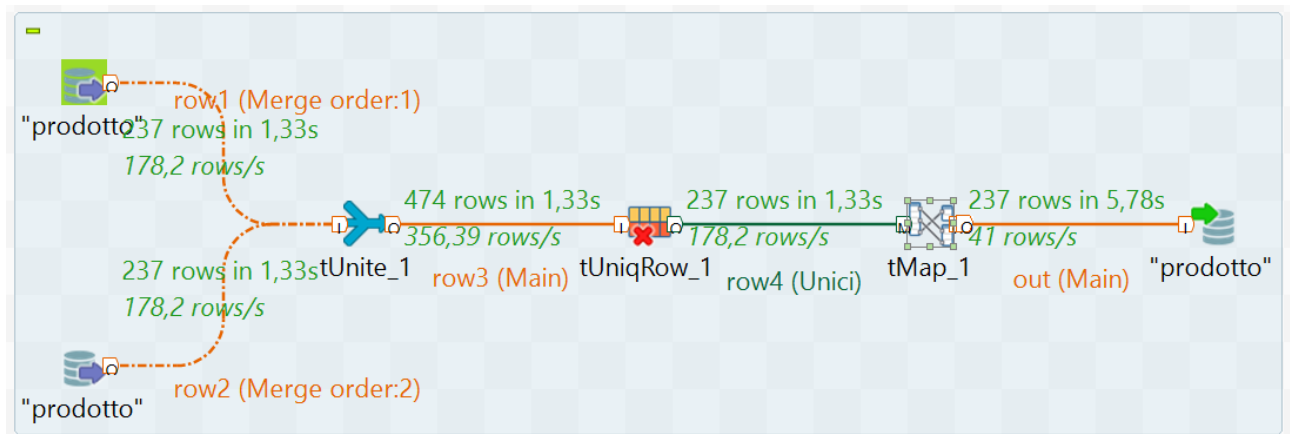
Magazzino



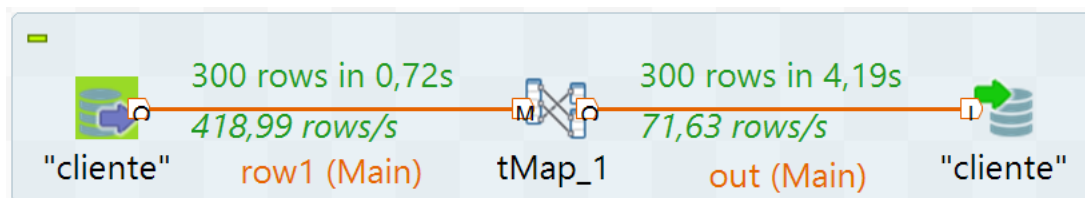
MateriaPrima



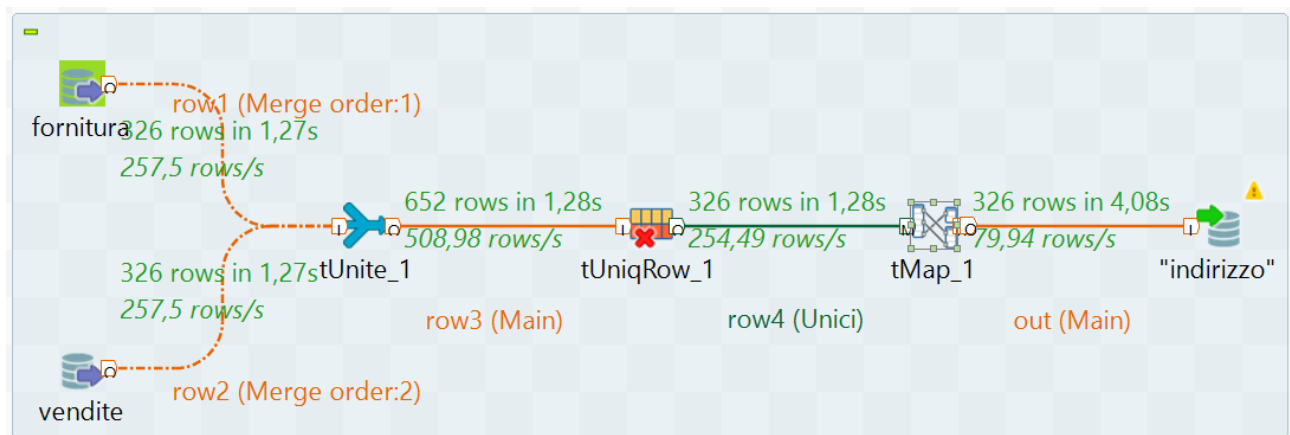
Prodotto



Cliente



Indirizzo



3.1.4 Progettazione concettuale

Nella fase di progettazione concettuale, i requisiti utente precedentemente raccolti vengono utilizzati per ottenere uno schema concettuale per il data mart, tramite il modello del *Dimensional Fact Model*.

Questa sezione contiene:

- la definizione dei fatti di interesse,
- la descrizione della costruzione e della modifica dell'albero degli attributi di ogni fatto,
- la descrizione delle misure e degli attributi
- gli schemi di fatto.

La fase di progettazione concettuale è stata svolta con l'ausilio di WAND, uno strumento CASE che assiste nella strutturazione di un data mart; effettua la progettazione concettuale in modo semiautomatico partendo dallo schema logico del DB.

Schema logico del DB riconciliato

```
cliente(ID#K%INT, Nome%ALFA50,IndirizzoCliente%INT:indirizzo!ID,email%ALFA50,
NumeroTelefonico%INT,Sconto%REAL)
ordinecliente(ID#K%INT, DataConsegna%DATA, ClienteOrdine%INT: cliente!ID)
vendita(OrdineVendita#K%INT: ordinecliente!ID, ProdottoVendita#K%INT: prodotto!ID,
Quantita%INT, Importo%REAL, PrezzoUnitario%REAL,Data%DATA)
contenuto(MateriaContenuta#K%INT: materiaprima!ID, ProdottoContenuto#K%INT:
prodotto!ID)
fornitore(IVA#K%ALFA12,Nome%ALFA50,email%ALFA50,IndirizzoFornitore%INT:indiri
zzo!ID ,NumeroTelefonico%INT)
fornitura(OrdineFornitura#K%INT: ordinefornitore!ID, MateriaPrimaFornitura#K%INT:
materiaprima!ID, Quantita%INT, Importo%REAL,Data%DATA, PrezzoUnitario%REAL)
magazzino(ID#K%INT, ProdottoMagazzino%INT: prodotto!ID, Lotto%ALFA50,
Prezzo%REAL, Nome%ALFA100, DataProduzione%DATA, DataScadenza%DATA)
materiaprima(ID#K%INT, Nome%ALFA100,DataScadenza%DATA,Lotto%ALFA50,
Tipologia%ALFA50)
ordinefornitore(ID#K%INT, IvaFornitore%ALFA12:fornitore!IVA)
indirizzo(ID#K%INT, Via%ALFA50, Numero%INT, Citta%ALFA50, Regione%ALFA50,
Stato%ALFA50)
prodotto(ID#K%INT,Tipologia%ALFA50 , Nome%ALFA100)
```

I fatti che si sono rivelati di interesse nella fase di analisi sono FORNITURA e VENDITA, corrispondenti alle relazioni Fornitura e Vendita.

L'albero degli attributi di FORNITURA, ottenuto tramite l'applicazione dell'algoritmo, è riportato di seguito.

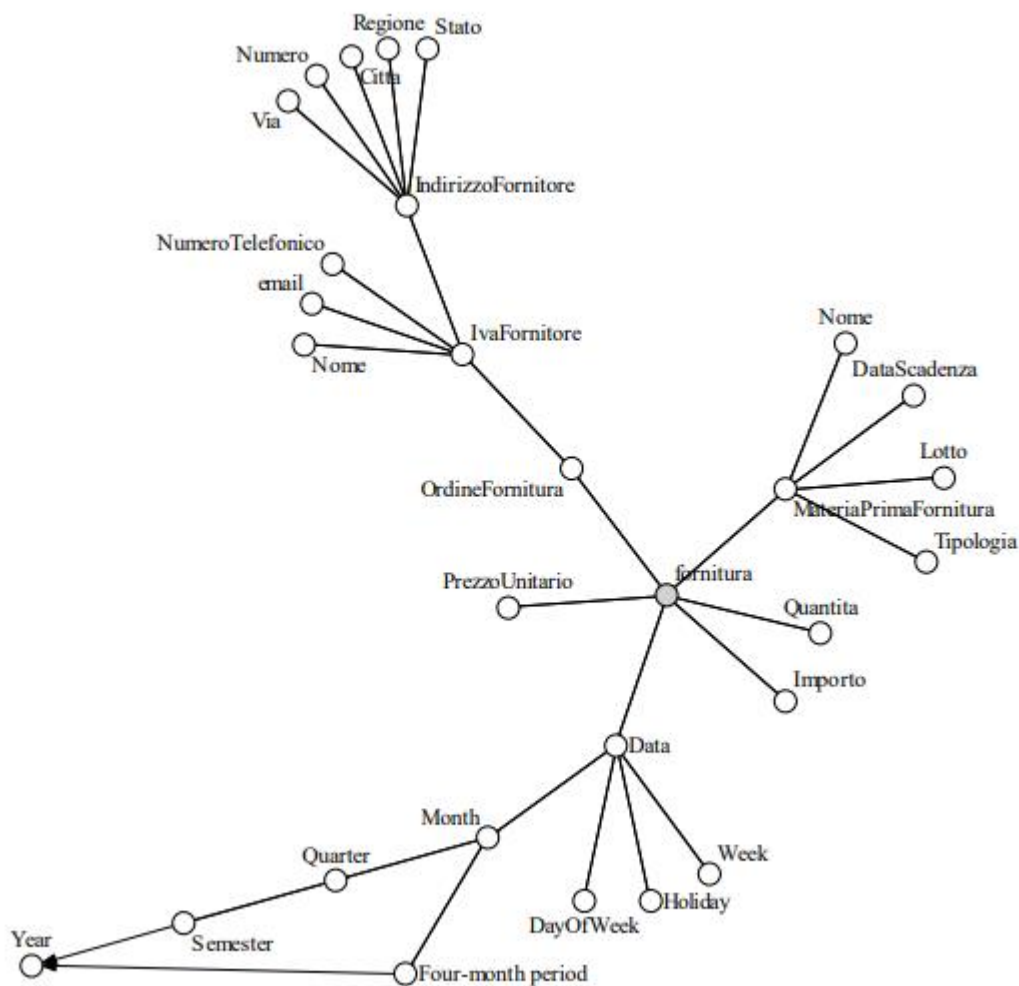


Figura 3.7 Albero degli attributi del fatto FORNITURA

Sono successivamente state fatte delle operazioni di riaggiustamento basate sulle necessità specifiche dell'impresa. In dettaglio:

- Si è scelto di focalizzare la dimensione temporale esclusivamente sulla data, il mese, il semestre e l'anno. Di conseguenza, sono state effettuate operazioni di potatura e innesto per eliminare alcuni livelli di dettaglio inseriti automaticamente dall'algoritmo.
- La struttura è stata radicalmente modificata sostituendo il progenitore dei nodi 'Regione' e 'Stato', con l'aggiunta corrispondente della dipendenza funzionale tra 'Città', 'Regione' e 'Stato'.
- Infine, per migliorare la chiarezza, sono stati apportati cambiamenti ai nomi di alcuni attributi.

Si ottiene, quindi, l'albero degli attributi modificato.

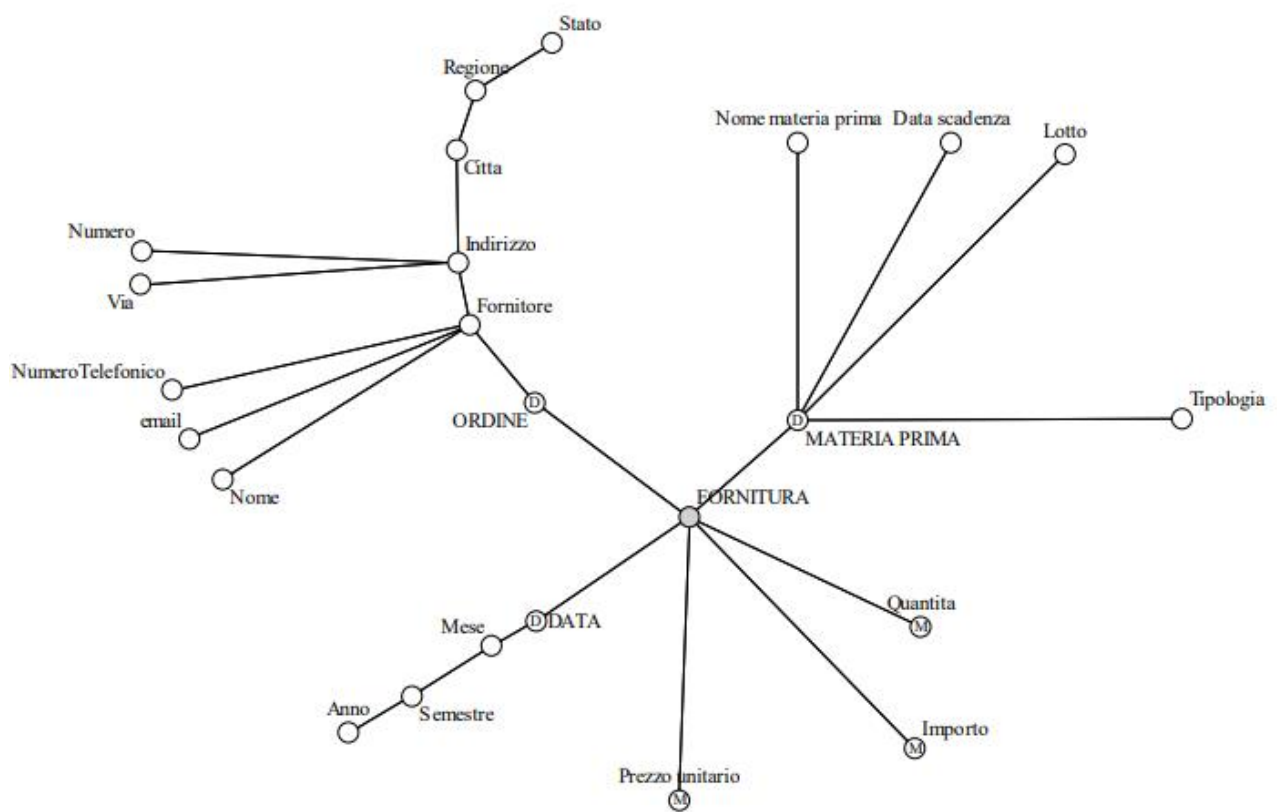


Figura 3.8 Albero degli attributi modificato del fatto FORNITURA

Ne risulta il seguente schema di fatto.

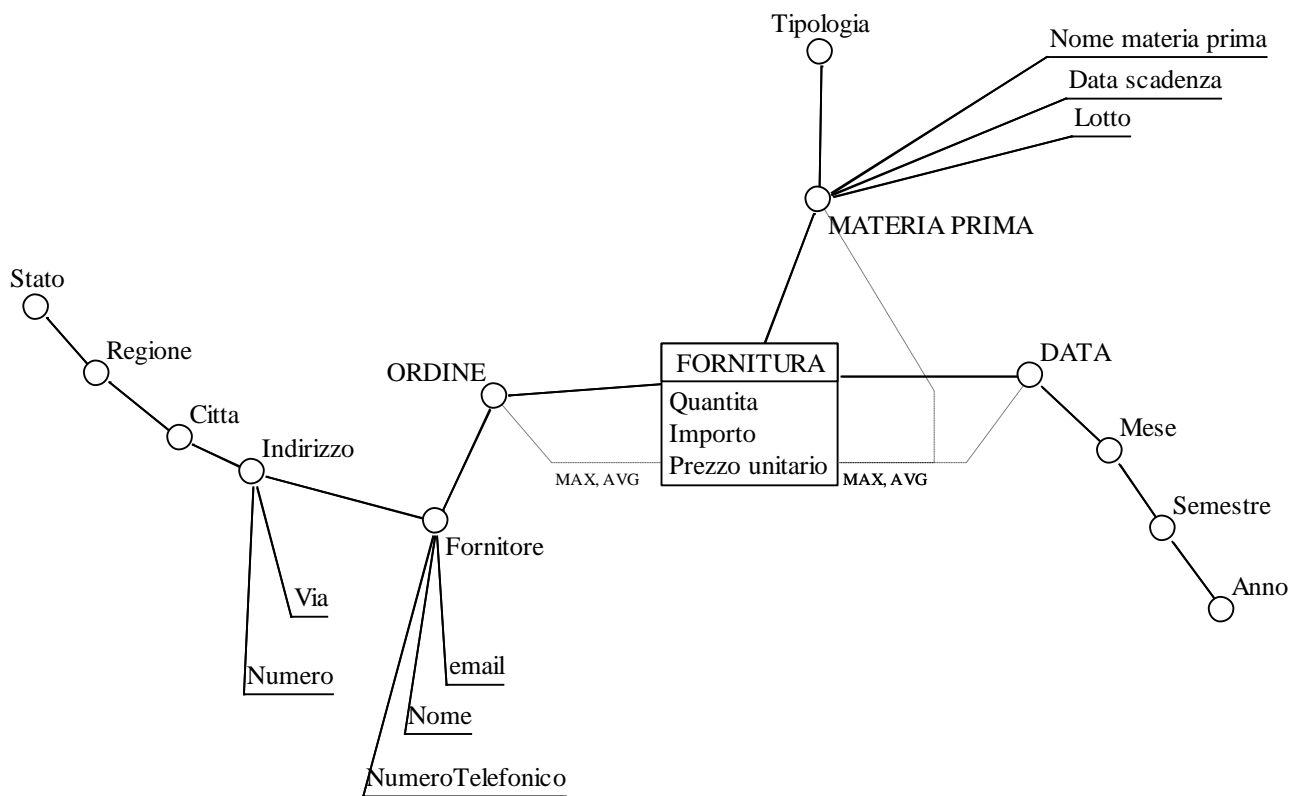


Figura 3.9 Schema di fatto di FORNITURA

Tabella 3.6 Glossario degli attributi del fatto FORNITURA

Name	Description	Domain	Rel. Info
<i>DATA</i>		Date	Data:fornitura (Date)
DATA:Mese		Date	
DATA:Semestre		Date	
DATA:Anno		Date	
<i>ORDINE</i>		Integer	ID:ordinefornitore (Integer)
ORDINE:Fornitore		String12	IVA:fornitore (String12)
ORDINE:Nome		String50	Nome:fornitore (String50)
ORDINE:email		String50	email:fornitore (String50)
ORDINE:NumeroTelefonico		Integer	NumeroTelefonico:fornitore (Integer)
ORDINE:Indirizzo		Integer	ID:indirizzo (Integer)
ORDINE:Via		String50	Via:indirizzo (String50)

ORDINE:Numero		Integer	Numero:indirizzo (Integer)
ORDINE:Citta		String50	Citta:indirizzo (String50)
ORDINE:Regione		String50	Regione:indirizzo (String50)
ORDINE:Stato		String50	Stato:indirizzo (String50)
MATERIA PRIMA		Integer	ID:materiaprima (Integer)
MATERIA PRIMA:Nome materia prima		String100	Nome:materiaprima (String100)
MATERIA PRIMA:Data scadenza		Date	DataScadenza:materiaprima (Date)
MATERIA PRIMA:Lotto		String50	Lotto:materiaprima (String50)
MATERIA PRIMA:Tipologia		String50	Tipologia:materiaprima (String50)

Tabella 3.7 Glossario delle misure del fatto FORNITURA

<i>Name</i>	<i>Description</i>	<i>Type</i>
Quantità	Quantità di materia prima richiesta nell'ordine.	INT
Importo	Importo complessivo della fornitura	DOUBLE
PrezzoUnitario	Prezzo unitario della materia prima	DOUBLE

La seguente figura mostra, invece, l'albero degli attributi del fatto VENDITA.

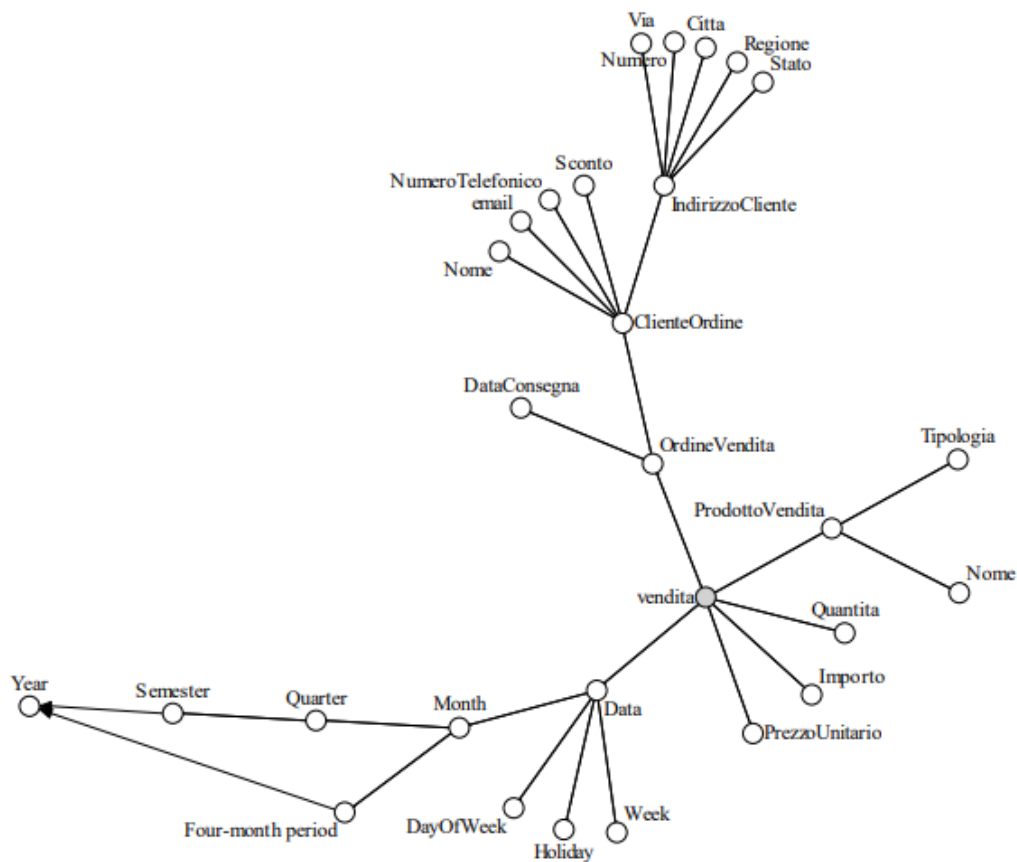


Figura 3.10 Albero degli attributi del fatto VENDITA

Anche su questo albero, generato automaticamente attraverso l'algoritmo, sono state eseguite le stesse operazioni di riaggiustamento effettuate sull'albero precedente.

Si ottiene, allora, l'albero degli attributi modificato.

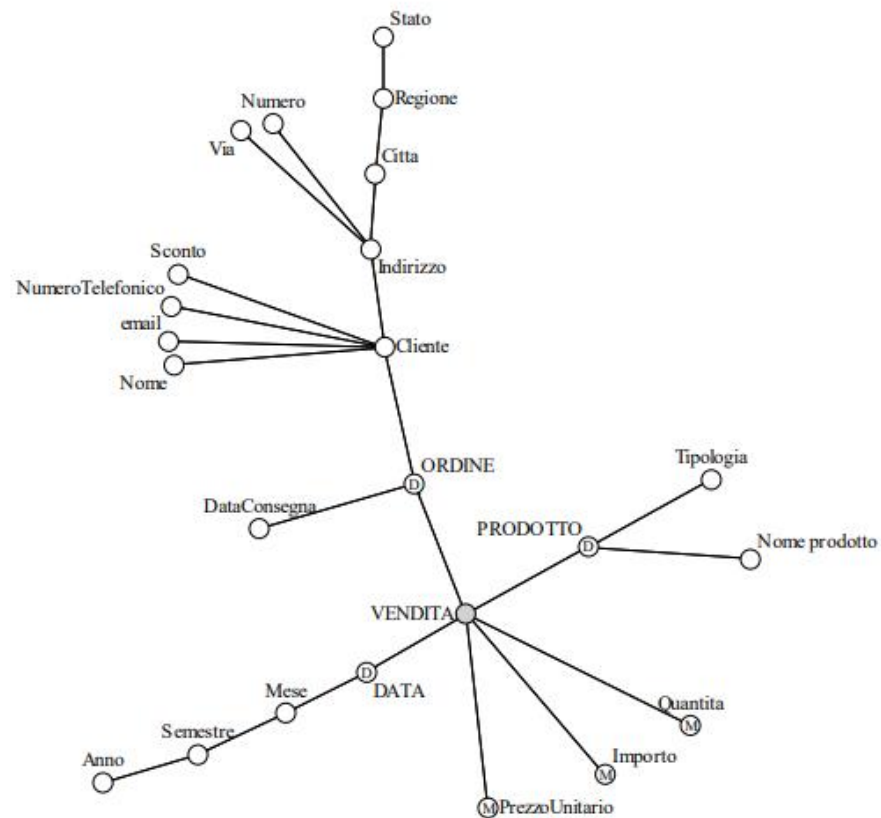


Figura 3.11 Albero degli attributi modificato del fatto VENDITA

Ne risulta il seguente schema di fatto.

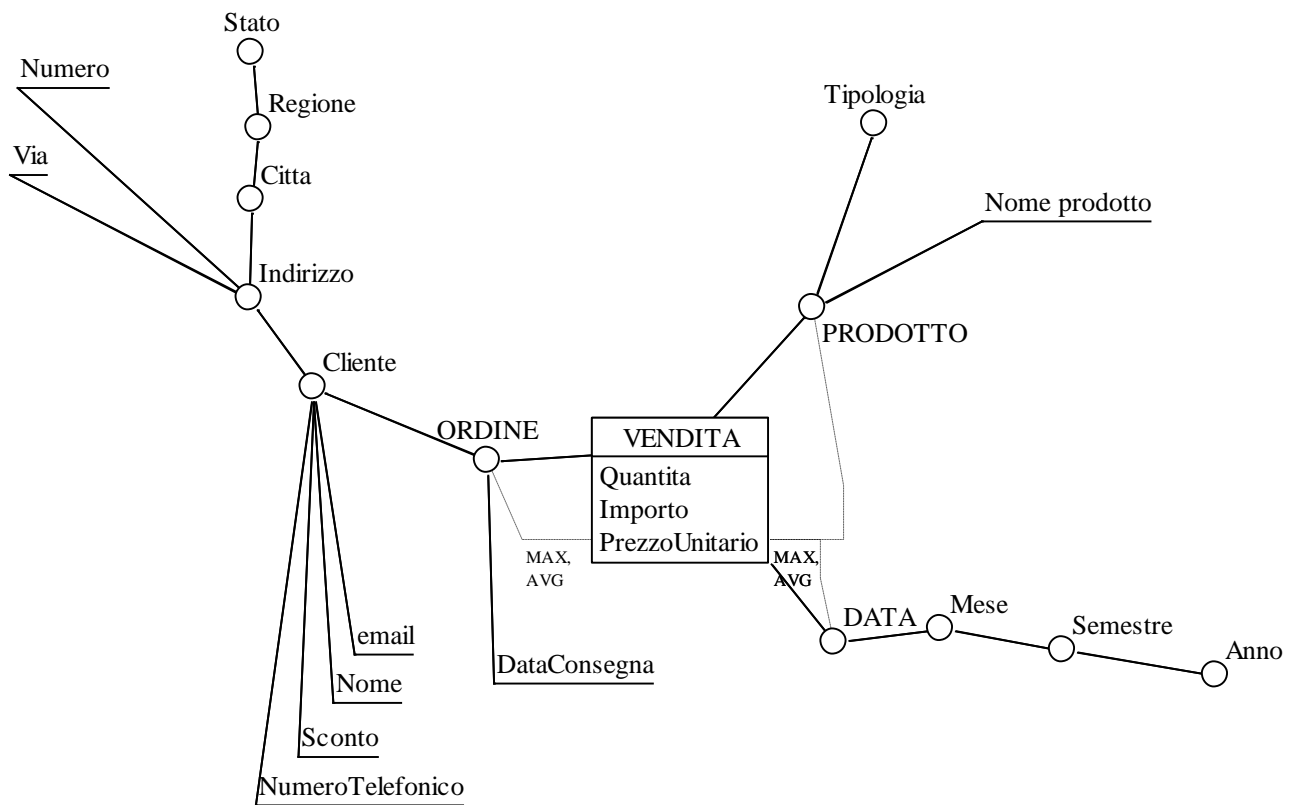


Figura 3.12 Schema di fatto di VENDITA

Tabella 3.8 Glossario degli attributi del fatto Vendita

Name	Description	Domain	Rel. Info
DATA		Date	Data:vendita (Date)
DATA:Mese		Date	
DATA:Semestre		Date	
DATA:Anno		Date	
ORDINE		Integer	ID:ordinecliente (Integer)
ORDINE:DataConsegna		Date	DataConsegna:ordinecliente (Date)
ORDINE:Cliente		Integer	ID:cliente (Integer)
ORDINE:Nome		String50	Nome:cliente (String50)
ORDINE:email		String50	email:cliente (String50)
ORDINE:NumeroTelefonico		Integer	NumeroTelefonico:cliente (Integer)
ORDINE:Sconto		Float	Sconto:cliente (Float)
ORDINE:Indirizzo		Integer	ID:indirizzo (Integer)

ORDINE:Via		String50	Via:indirizzo (String50)
ORDINE:Numero		Integer	Numero:indirizzo (Integer)
ORDINE:Citta		String50	Citta:indirizzo (String50)
ORDINE:Regione		String50	Regione:indirizzo (String50)
ORDINE:Stato		String50	Stato:indirizzo (String50)
<i>PRODOTTO</i>		Integer	ID:prodotto (Integer)
PRODOTTO:Tipologia		String50	Tipologia:prodotto (String50)
PRODOTTO:Nome prodotto		String100	Nome:prodotto (String100)

Tabella 3.9 Glossario delle misure del fatto Vendita

<i>Name</i>	<i>Description</i>	<i>Type</i>
Quantità	Quantità di prodotto richiesta	INT
Importo	Importo complessivo	DOUBLE
PrezzoUnitario	Prezzo unitario del prodotto	DOUBLE

Evidentemente, sia per il fatto Vendita che Fornitura, le misure 'Incasso' e 'Quantità' risultano essere additive, indicando una somma cumulativa. Per quanto riguarda la misura 'PrezzoUnitario', verrà aggregata utilizzando sia l'operatore MAX che l'operatore AVG. Pertanto, rappresenterà sia il prezzo unitario massimo che il prezzo unitario medio per la materia prima e il prodotto.

3.1.5 Progettazione logica

In questa fase, si produce lo schema logico del data mart.

Prima di tutto, è necessario scegliere tra un'implementazione ROLAP (su DBMS relazionale) e una MOLAP (su DBMS multidimensionale).

Il principale svantaggio dell'implementazione è ROLAP riguarda le prestazioni, perché le interrogazioni richiedono costose operazioni di join. Un sistema MOLAP, invece, si basa su un modello logico pensato per i dati multidimensionali, che vengono fisicamente memorizzati in *array*.

Tuttavia, non esistendo ancora uno standard per il modello logico multidimensionale, le diverse implementazioni MOLAP hanno pochi punti in comune.

Data la limitata diffusione sul mercato dei sistemi MOLAP, per questo progetto si è scelto di utilizzare un'implementazione ROLAP.

Lo schema logico di base è lo *schema a stella*, modello in grado di conciliare i concetti di fatto, dimensione di analisi, gerarchie e misure con le ben note strutture di relazioni, attributi e vincoli di integrità.

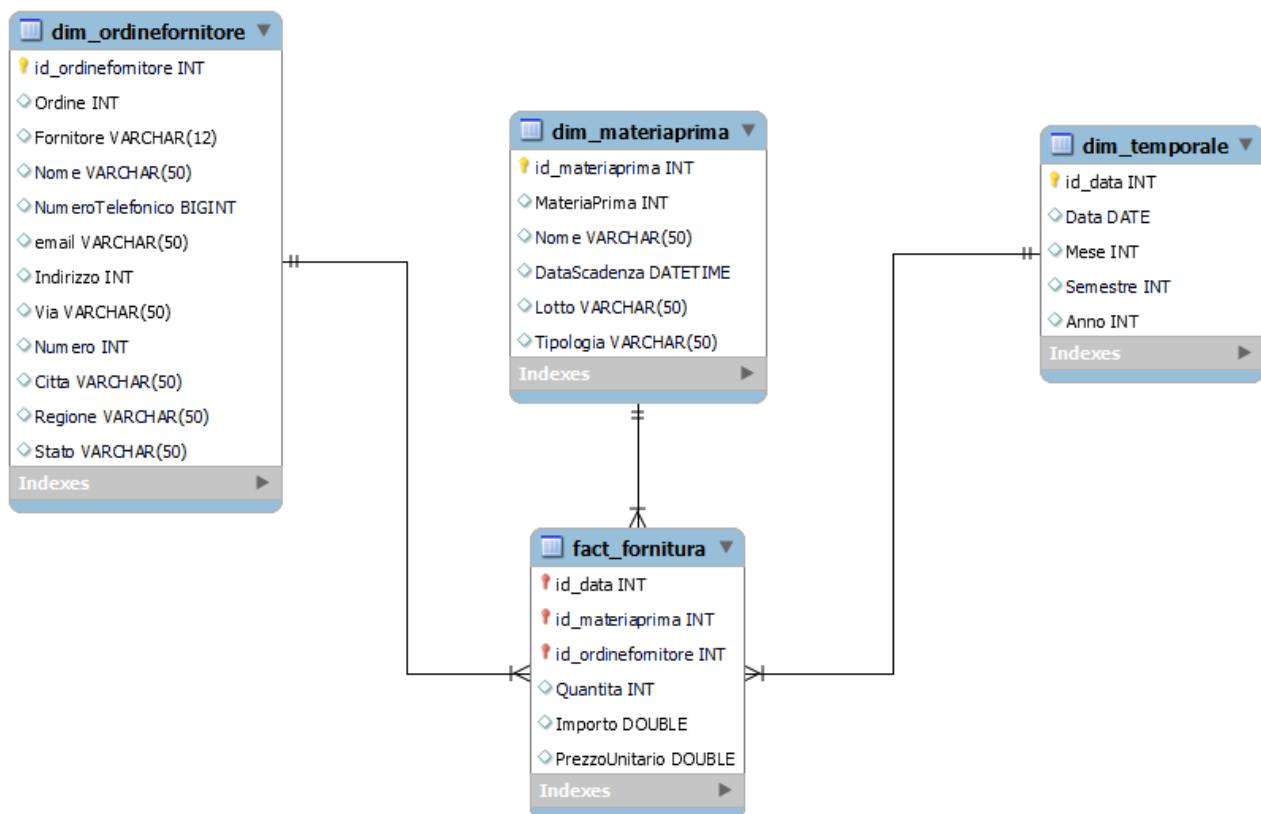


Figura 3.13 Schema concettuale

Schema logico

dim_ordinefornitore(id_ordinefornitore, Ordine, Fornitore, Nome, Numero Telefonico, email, Indirizzo, Via, Numero, Citta, Regione, Stato)

dim_materiaprima(id_materiaprima, MateriaPrima, Nome, DataScadenza, Lotto, Tipologia)

dim_temporale(id_data, Data, Mese, Semestre, Anno)

fact_fornitura(id_data: **dim_temporale**, id_materiaprima: **dim_materiaprima**, id_ordinefornitore: **dim_ordinefornitore**, Quantita, Importo, PrezzoUnitario)

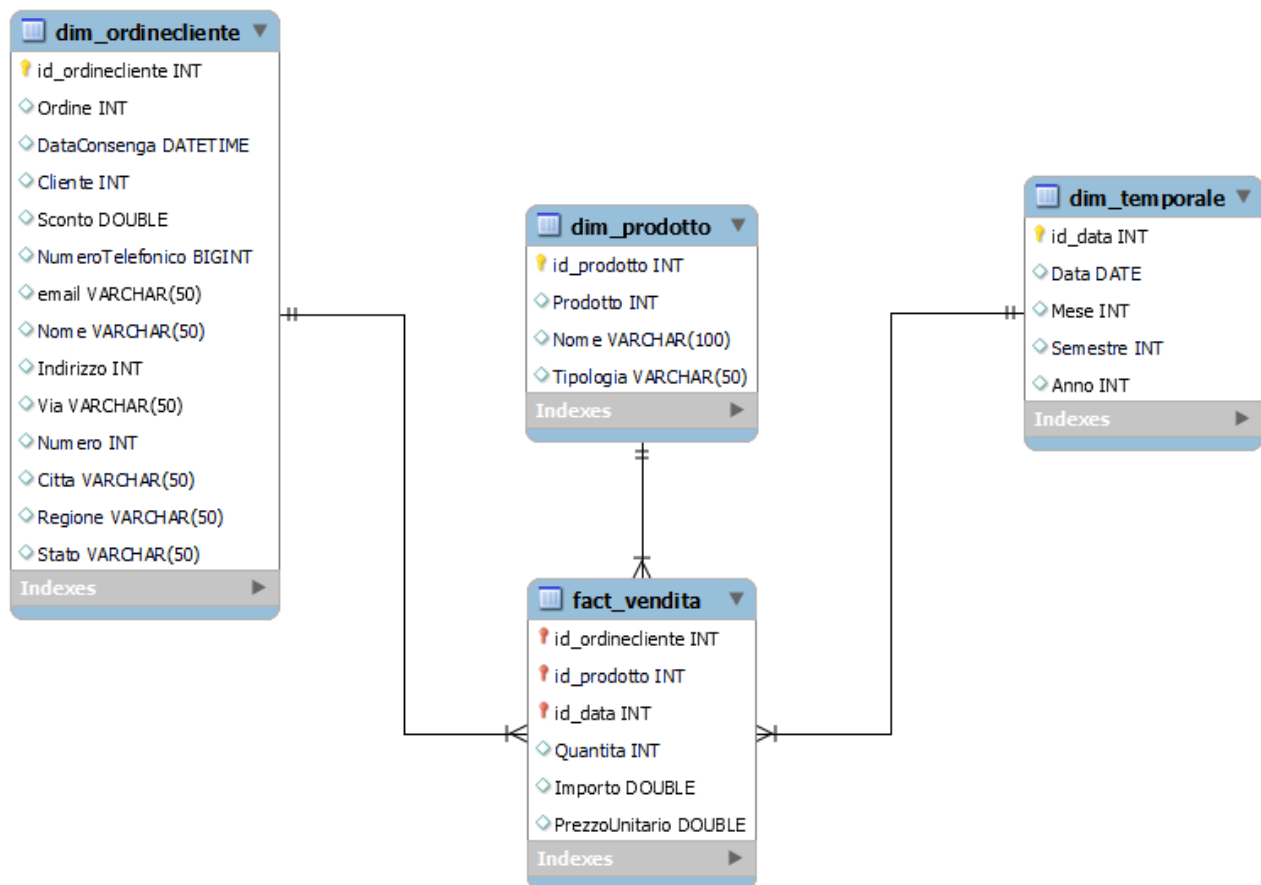


Figura 3.14 Schema concettuale

Schema logico

dim_ordinecliente(id_ordinecliente, Ordine, DataConsegna, Cliente, Sconto, Numero Telefonico, email, Nome, Indirizzo, Via, Numero, Citta, Regione, Stato)

dim_prodotto(id_prodotto, Prodotto, Nome, Tipologia)

dim_temporale(id_data, Data, Mese, Semestre, Anno)

fact_fornitura(id_ordinecliente: **dim_ordinecliente**, id_prodotto: **dim_prodotto**, id_data: **dim_temporale**, Quantita, Importo, PrezzoUnitario)

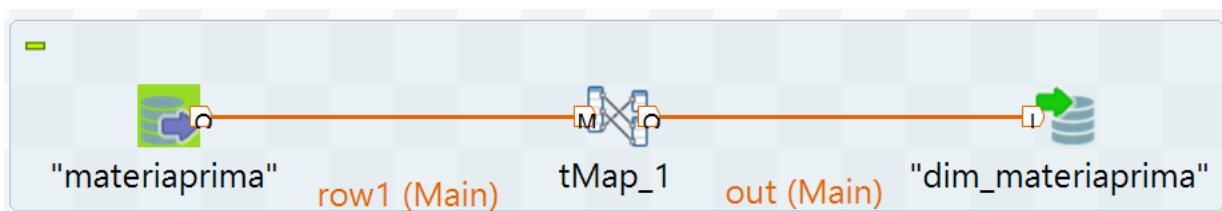
3.1.6 Progettazione dell'alimentazione: dal riconciliato al derivato

Durante questa fase, vengono sviluppate le procedure che consentono il confronto dei dati archiviati nel livello riconciliato con gli schemi a stella utilizzati nell'ambito multidimensionale. Oltre alle procedure per la denormalizzazione dei dati e l'inserimento delle chiavi surrogate, vengono definite anche quelle per il calcolo di dati derivati che non sono presenti a livello operativo.

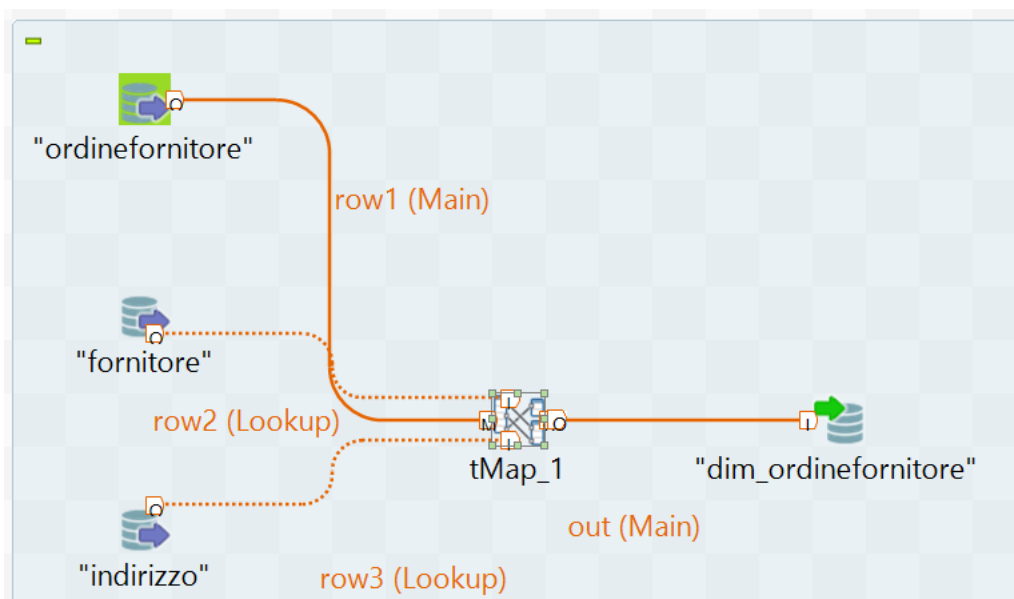
Come precedentemente realizzato, si fa uso del software Talend per popolare il data mart a partire dai dati presenti nel database riconciliato.

Di seguito, sono riportati i *job* utilizzati a tal fine.

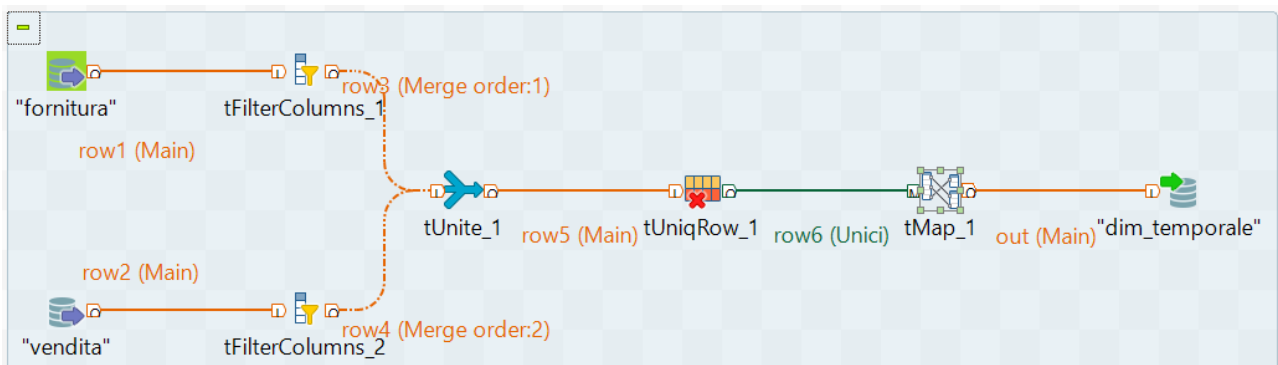
Dimensione Materia Prima



Dimensione Ordine Fornitore



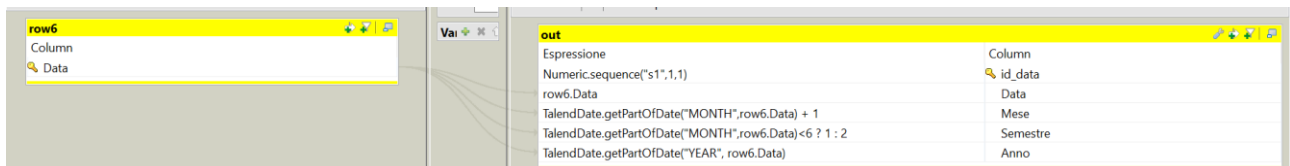
Dimensione temporale



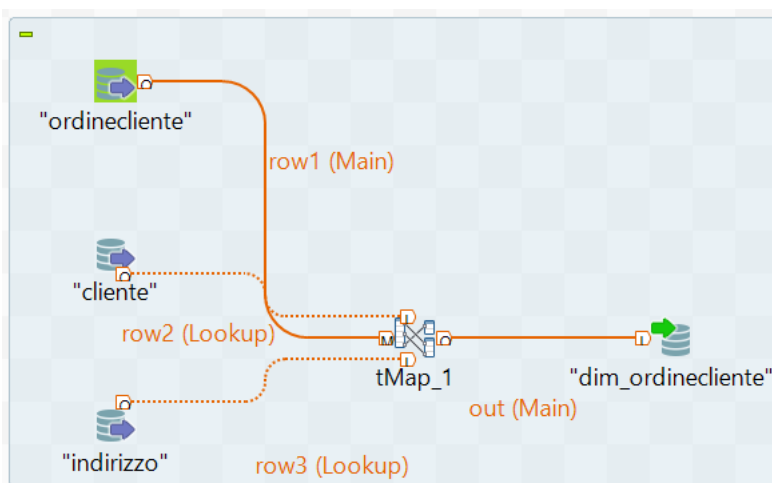
Dall'analisi dei requisiti utente, è emersa la necessità di aggiungere dei livelli di dettaglio nella gerarchia temporale.

Di conseguenza, sono stati introdotti nella tabella relativa alla dimensione temporale, coerentemente con lo schema di fatto, alcuni attributi mediante l'utilizzo della funzione *TalendDate.getPartOfDate()*:

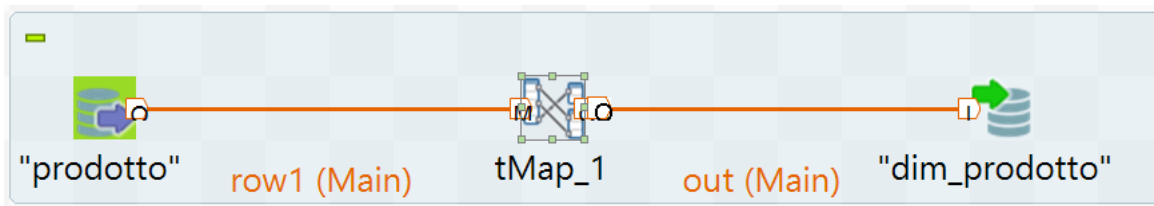
- Mese: è ottenuto attraverso l'applicazione diretta della funzione, specificando la parte della data desiderata come parametro. Il risultato viene incrementato di uno poiché la funzione restituisce i mesi nell'intervallo [0, 11], dove gennaio è rappresentato dal numero 0.
- Semestre: viene derivato applicando la funzione come descritto in precedenza. Se il mese è inferiore o uguale a 6, la funzione restituisce 1; in caso contrario, restituisce 2.
- Anno: è ottenuto anch'esso tramite l'applicazione della funzione, specificando 'YEAR' come parametro.



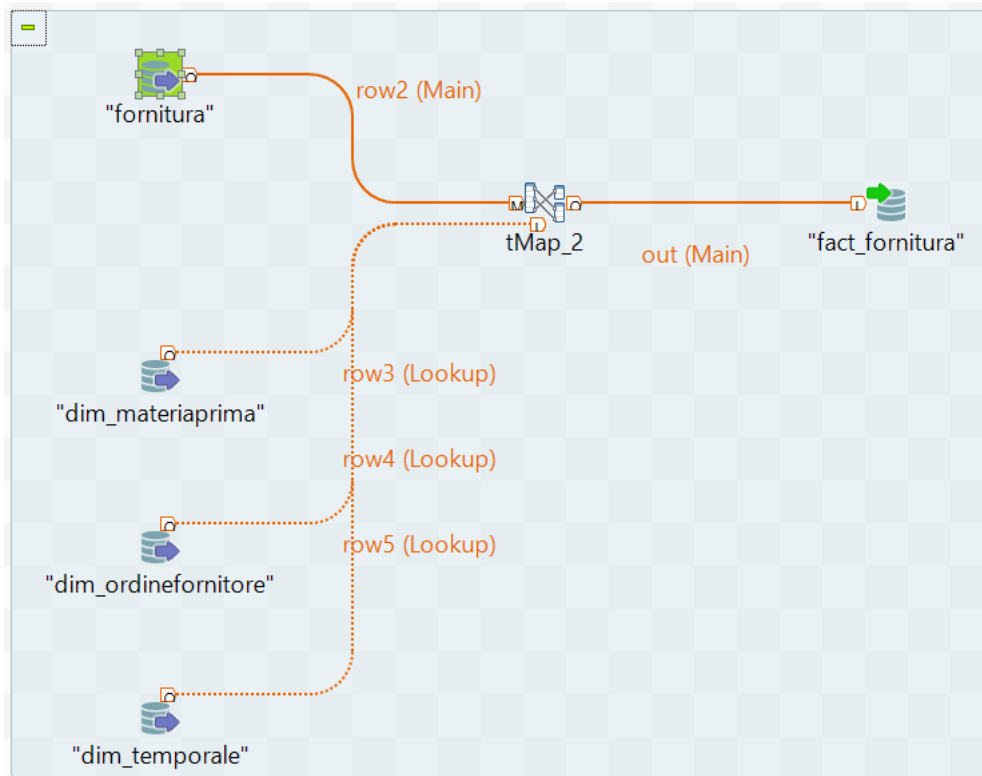
Dimensione Ordine Cliente



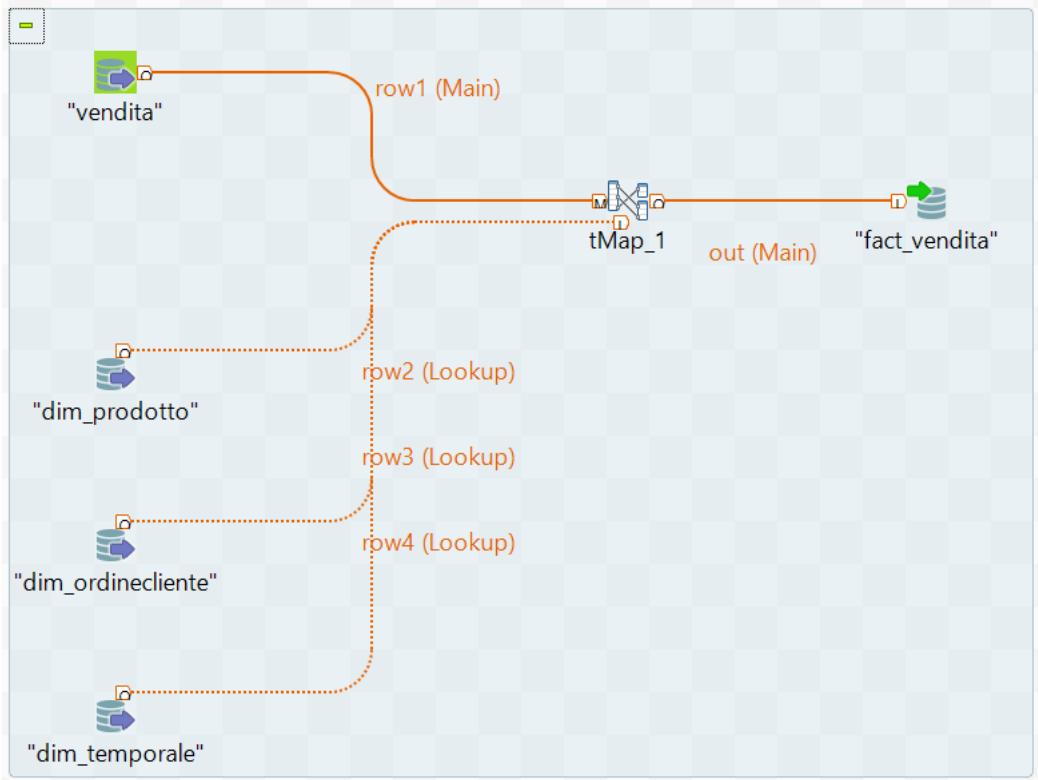
Dimensione Prodotto



Fatto Fornitura



Fatto Vendita



3.1.7 Cubi OLAP

Le interrogazioni OLAP sono di diretto interesse dell'utente finale.

Un'interrogazione OLAP consiste in una successione di interrogazioni, ognuna delle quali, eventualmente, è suggerita all'utente dal risultato della precedente. Un'analisi di questo tipo fornisce all'utente un'esperienza interattiva. Ogni interrogazione si traduce nell'applicazione di un operatore OLAP ai dati (*roll-up, drill-down, slicing, ecc.*).

I cubi OLAP sono stati realizzati tramite il software *Mondrian Schema Workbench*.

I risultati ottenuti durante la progettazione concettuale sono stati utilizzati per individuare dimensioni, gerarchie e misure per ogni fatto di interesse.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<Schema name="Schema">
  <Cube name="Cubo Vendite" visible="true" cache="true" enabled="true">
    <Table name="fact_vendita"> </Table>
    <Dimension type="StandardDimension" visible="true" foreignKey="id_data" highCardinality="false"
      name="dim_temporale" caption="Dimensione temporale">
      <Hierarchy name="Tempo" visible="true" hasAll="true" allMemberName="All" primaryKey="id_data">
        <Table name="dim_temporale"> </Table>
        <Level name="Anno" visible="true" column="Anno" type="Integer" internalType="int"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Anno"> </Level>
        <Level name="Semestre" visible="true" column="Semestre" type="Integer" internalType="int"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Semestre">
          <NameExpression>
            <SQL dialect="generic">
              <![CDATA[ CASE WHEN Semestre = 1 THEN "Primo semestre" WHEN Semestre= 2 THEN "Secondo semestre"
                END ]]>
            </SQL>
          </NameExpression>
        </Level>
        <Level name="Mese" visible="true" column="Mese" type="Integer" internalType="int"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Mese">
          <NameExpression>
            <SQL dialect="generic">
              <![CDATA[ CASE WHEN Mese = 1 THEN "Gennaio" WHEN Mese= 2 THEN "Febbraio" WHEN Mese= 3 THEN
                "Marzo" WHEN Mese= 4 THEN "Aprile" WHEN Mese= 5 THEN "Maggio" WHEN Mese= 6 THEN "Giugno" WHEN
                Mese= 7 THEN "Luglio" WHEN Mese= 8 THEN "Agosto" WHEN Mese= 9 THEN "Settembre" WHEN Mese= 10
                THEN "Ottobre" WHEN Mese= 11 THEN "Novembre" WHEN Mese= 12 THEN "Dicembre" END ]]>
            </SQL>
          </NameExpression>
        </Level>
        <Level name="Data" visible="true" column="Data" type="Date" internalType="Object"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Data"> </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" visible="true" foreignKey="id_prodotto" highCardinality="false"
      name="dim_prodotto" caption="Dimensione prodotto">
      <Hierarchy name="Prodotto" visible="true" hasAll="true" allMemberName="All" primaryKey="id_prodotto">
        <Table name="dim_prodotto"> </Table>
        <Level name="Tipologia" visible="true" column="Tipologia" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Tipologia"> </Level>
        <Level name="Prodotto" visible="true" column="Nome" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Prodotto"> </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" visible="true" foreignKey="id_ordinecliente" highCardinality="false"
      name="dim_ordinecliente" caption="Dimensione ordine cliente">
      <Hierarchy name="Ordine Cliente" visible="true" hasAll="true" allMemberName="All">
        <Table name="dim_ordinecliente"> </Table>
        <Level name="Stato" visible="true" column="Stato" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Stato"> </Level>
        <Level name="Regione" visible="true" column="Regione" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Regione"> </Level>
        <Level name="Città" visible="true" column="Città" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Città"> </Level>
        <Level name="Indirizzo" visible="true" column="Via" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Indirizzo cliente"> </Level>
        <Level name="Cliente" visible="true" column="Nome" type="String" internalType="String"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Cliente"> </Level>
        <Level name="Ordine cliente" visible="true" column="Ordine" type="Integer" internalType="int"
          uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Ordine cliente"> </Level>
      </Hierarchy>
    </Dimension>
    <Measure name="Importo" column="Importo" datatype="Numeric" aggregator="sum" caption="Importo"
      visible="true"> </Measure>
    <Measure name="Quantita" column="Quantita" datatype="Integer" aggregator="sum" caption="Quantità"
      visible="true"> </Measure>
    <Measure name="Prezzo unitario massimo" column="PrezzoUnitario" datatype="Numeric" aggregator="max"
      caption="Prezzo unitario massimo" visible="true"> </Measure>
    <Measure name="Prezzo unitario medio" column="PrezzoUnitario" datatype="Numeric" aggregator="avg"
      caption="Prezzo unitario medio" visible="true"> </Measure>
  </Cube>
  <Cube name="Cubo Fornitura" visible="true" cache="true" enabled="true">
```

Figura 3.15


```

<Table name="fact_fornitura"> </Table>
<Dimension type="StandardDimension" visible="true" foreignKey="id_data" highCardinality="false"
name="dim_temporale" caption="Dimensione temporale">
  <Hierarchy name="Tempo" visible="true" hasAll="true" allMemberName="All" primaryKey="id_data">
    <Table name="dim_temporale"> </Table>
    <Level name="Anno" visible="true" column="Anno" type="Integer" internalType="int"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Anno"> </Level>
    <Level name="Semestre" visible="true" column="Semestre" type="Integer" internalType="int"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Semestre">
      <NameExpression>
        <SQL dialect="generic">
          <![CDATA[ CASE WHEN Semestre = 1 THEN "Primo semestre" WHEN Semestre= 2 THEN "Secondo semestre"
END ]]>
        </SQL>
      </NameExpression>
    </Level>
    <Level name="Mese" visible="true" column="Mese" type="Integer" internalType="int"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Mese">
      <NameExpression>
        <SQL dialect="generic">
          <![CDATA[ CASE WHEN Mese = 1 THEN "Gennaio" WHEN Mese= 2 THEN "Febbraio" WHEN Mese= 3 THEN
"Marzo" WHEN Mese= 4 THEN "Aprile" WHEN Mese= 5 THEN "Maggio" WHEN Mese= 6 THEN "Giugno" WHEN
Mese= 7 THEN "Luglio" WHEN Mese= 8 THEN "Agosto" WHEN Mese= 9 THEN "Settembre" WHEN Mese= 10
THEN "Ottobre" WHEN Mese= 11 THEN "Novembre" WHEN Mese= 12 THEN "Dicembre" END ]]>
        </SQL>
      </NameExpression>
    </Level>
    <Level name="Data" visible="true" column="Data" type="Date" internalType="Object"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Data"> </Level>
  </Hierarchy>
</Dimension>
<Dimension type="StandardDimension" visible="true" foreignKey="id_materiaprima" highCardinality="false"
name="dim_materiaprima" caption="Materia prima">
  <Hierarchy name="Materia prima" visible="true" hasAll="true" allMemberName="All"
primaryKey="id_materiaprima">
    <Table name="dim_materiaprima"> </Table>
    <Level name="Tipologia" visible="true" column="Tipologia" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never"> </Level>
    <Level name="Materia prima" visible="true" column="Nome" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Materia prima"> </Level>
  </Hierarchy>
</Dimension>
<Dimension type="StandardDimension" visible="true" foreignKey="id_ordinefornitore"
highCardinality="false" name="dim_ordinefornitore" caption="Dimensione ordine fornitore">
  <Hierarchy name="Ordine fornitore" visible="true" hasAll="true" allMemberName="All"
primaryKey="id_ordinefornitore">
    <Table name="dim_ordinefornitore"> </Table>
    <Level name="Stato" visible="true" column="Stato" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Stato"> </Level>
    <Level name="Regione" visible="true" column="Regione" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Regione"> </Level>
    <Level name="Città" visible="true" column="Città" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Città"> </Level>
    <Level name="Indirizzo" visible="true" column="Via" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Indirizzo"> </Level>
    <Level name="Fornitore" visible="true" column="Nome" type="String" internalType="String"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Fornitore"> </Level>
    <Level name="Ordine fornitore" visible="true" column="Ordine" type="Integer" internalType="int"
uniqueMembers="false" levelType="Regular" hideMemberIf="Never" caption="Ordine"> </Level>
  </Hierarchy>
</Dimension>
<Measure name="Importo" column="Importo" datatype="Numeric" aggregator="sum" caption="Importo"
visible="true"> </Measure>
<Measure name="Quantità" column="Quantità" datatype="Integer" aggregator="sum" caption="Quantità"
visible="true"> </Measure>
<Measure name="Prezzo unitario massimo" column="PrezzoUnitario" datatype="Numeric" aggregator="max"
caption="Prezzo unitario massimo" visible="true"> </Measure>
<Measure name="Prezzo unitario medio" column="PrezzoUnitario" datatype="Numeric" aggregator="avg"
caption="Prezzo unitario medio" visible="true"> </Measure>
</Cube>
</Schema>

```

Figura 3.16



Figura 3.17 Cubo Fornitura



Figura 3.18 Cubo Vendita

Come evidenziato dallo schema, sono state implementate espressioni di nome (Name Expression) nei livelli 'Mese' e 'Semestre' della gerarchia temporale. Ciò consente la visualizzazione del nome del mese e del semestre, migliorando la chiarezza e la leggibilità nelle rappresentazioni visive dei dati, anziché utilizzare i numeri corrispondenti.

SQL	
Attribute	Value
cdata	<pre>CASE WHEN Mese = 1 THEN "Gennaio" WHEN Mese= 2 THEN "Febbraio" WHEN Mese= 3 THEN "Marzo" WHEN Mese= 4 THEN "Aprile" WHEN Mese= 5 THEN "Maggio" WHEN Mese= 6 THEN "Giugno" WHEN Mese= 7 THEN "Luglio" WHEN Mese= 8 THEN "Agosto" WHEN Mese= 9 THEN "Settembre" WHEN Mese= 10 THEN "Ottobre" WHEN Mese= 11 THEN "Novembre" WHEN Mese= 12 THEN "Dicembre" END</pre>

Figura 3.19

SQL	
Attribute	Value
cdata	<pre>CASE WHEN Semestre = 1 THEN "Primo semestre" WHEN Semestre= 2 THEN "Secondo semestre" END</pre>

Figura 3.20

Drag the member here if you want to use it as a slicer

Measures						
				Importo	Quantità	Prezzo unitario massimo
						Prezzo unitario medio
All	All	All		74,981	29,562	3.5
2022	All	All		26,693.5	10,421	3.5
2023	All	All		48,287.5	19,141	3.5
Primo semestre	All	All		28,975.5	11,451	3.5
Secondo semestre	All	All		19,312	7,690	3.5
	Pasta	All		3,887	1,592	3.5
		Brasile		64	30	3.5
		Francia		89.5	35	3.5
		Germania		10.5	7	1.5
		Giappone				
		Italia		3,637.5	1,487	3.5
		Abruzzo				
		Basilicata		144	60	3.5
		Calabria		17.5	7	2.5
		Campania		300	122	3.5
		Emilia-Romagna		489.5	191	3.5
		Friuli-Venezia Giulia		166	76	3.5
		Lazio		375.5	149	3.5
		Liguria		121	40	3.5
		Lombardia		396	170	3.5
		Marche		94	50	2.5
		Molise				
		Piemonte		122.5	35	3.5
		Puglia		350	128	3.5
		Sardegna		40.5	17	3.5
		Sicilia		111.5	55	3.5
		Toscana		353.5	145	3.5
		Trentino-Alto Adige		76.5	27	3.5
		Umbria		63.5	31	2.5
		Valle d'Aosta		22.5	15	1.5
		Veneto		393.5	169	3.5
		Paesi Bassi				
		Regno Unito				
		Russia		9	6	1.5
		Spagna		76.5	27	3.5
		Stati Uniti				
		Ungheria				
	Sott'olio	All		4,901.5	1,931	3.5
	Spezia	All		4,484	1,798	3.5
	Sugo	All		6,039.5	2,369	3.5

Figura 3.21 Vista relativa al fatto Fornitura

Drag the member here if you want to use it as a slicer

Measures			Importo	Quantità	Prezzo unitario massimo	Prezzo unitario medio
2022	All	All	450,096.707	66,966	10.5	7.017
	Pasta	All	156,427.609	23,514	10.5	6.967
		All	42,820.89	6,402	10	7.037
		Australia	176.4	18	9.8	9.8
		Germania	96.3	26	5.5	4.5
		Giappone				
		Italia	41,064.415	6,134	10	7.053
		Abruzzo	267	39	7.5	7.5
		Basilicata	1,836.67	267	9.5	7.2
		Calabria	2,048.32	308	10	7.069
		Campania	2,132.636	297	9.8	7.318
		Emilia-Romagna	2,666.632	411	10	6.805
		Friuli-Venezia Giulia	1,153.92	177	9.8	7.064
		Lazio	6,117.319	869	10	7.408
		Liguria	1,883.75	293	10	6.863
		Lombardia	6,161.84	923	9.8	6.977
		Marche	633.84	100	9.8	7.1
		Molise				
		Piemonte	2,182.944	306	10	7.239
		Puglia	4,125.563	635	9.9	6.942
		Sardegna	1,437.02	199	9.2	7.709
		Sicilia	1,339.79	216	9.8	6.55
		Toscana	2,381.769	367	10	6.905
		Trentino-Alto Adige	1,032.73	147	9.2	7.037
		Umbria	1,527.729	227	9.9	7.142
		Valle d'Aosta	322.02	44	9.2	7.675
		Veneto	1,812.923	309	9.2	6.171
		Norvegia				
		Regno Unito				
		Russia				
		Spagna	1,229.475	194	9.9	6.7
		Stati Uniti				
		Ungheria	254.3	30	9.2	7.65
2023	Sott'olio	All	42,483.196	6,511	10.1	6.864
	Spezia	All	39,703.64	5,864	10.1	7.039
	Sugo	All	31,419.883	4,737	10.5	6.927
	All	All	293,669.098	43,452	10.5	7.044

Figura 3.22 Vista relativa al fatto Vendita

3.1.8 Reportistica

Un altro approccio all'analisi dei dati disponibili in azienda sono i *report*. Nei casi in cui determinate interrogazioni, di forma fissa, siano necessarie con cadenza periodica, queste possono essere disegnate in fase di progettazione e incluse in applicazioni, in modo che possano essere autonomamente eseguite dagli utenti.

Di seguito, sono riportati due esempi di interrogazioni, con relativi report.

```
SELECT m.Nome, t.Data, o.Via, o.Citta,  
SUM(f.Importo) as Importo  
FROM dim_materiaprima m, dim_temporale t, dim_ordinefornitore o, fact_fornitura f  
WHERE m.id_materiaprima = f.id_materiaprima  
AND t.id_data=f.id_data  
AND o.id_ordinefornitore = f.id_ordinefornitore  
AND m.Tipologia= "Pasta"  
AND t.Anno= 2023  
AND t.Semestre=1  
AND o.Regione= "Lombardia"  
GROUP BY m.Nome, t.Data, o.Via  
order by t.Data;
```

Importo totale speso per acquistare le materie prime appartenenti alla tipologia 'Pasta' in Lombardia nel primo semestre del 2023:

Nome	Data	Via	Città	Importo
PEPE ROSA	Jan 1, 2023	Via dei Condotti	Milano	13.50 €
CANNELLA	Jan 10, 2023	Via della Moscova	Monza	31.50 €
CURRY	Jan 10, 2023	Via della Moscova	Monza	12.00 €
PEPERONCINO FRANTUMATO	Jan 12, 2023	Via della Repubblica	Pavia	17.50 €
CANNELLA IN POLVERE	Jan 12, 2023	Via della Repubblica	Pavia	10.50 €
MENTA	Jan 13, 2023	Via della Repubblica	Pavia	24.50 €
ERBA CIPOLLINA	Jan 13, 2023	Via della Repubblica	Pavia	24.50 €
CANNELLA IN POLVERE	Jan 13, 2023	Via della Repubblica	Pavia	12.00 €
SALE GROSSO	Jan 20, 2023	Corso Garibaldi	Bergamo	7.50 €
TIMO	Jan 20, 2023	Corso Garibaldi	Bergamo	20.00 €
MELISSA	Feb 7, 2023	Via dei Leoni	Lodi	25.00 €
ORIGANO	Feb 7, 2023	Via dei Leoni	Lodi	13.50 €
PREZZEMOLO	Feb 7, 2023	Via dei Leoni	Lodi	31.50 €
PEPERONCINO FRANTUMATO	Feb 16, 2023	Via della Moscova	Cremona	9.00 €
ARANCIA	Mar 10, 2023	Via Garibaldi	Milano	13.50 €
PREZZEMOLO	Mar 10, 2023	Corso Buenos Aires	Milano	35.00 €
TIMO	Mar 10, 2023	Via Garibaldi	Milano	20.00 €
PEPE ROSA	Mar 13, 2023	Via Dante Alighieri	Varese	12.00 €
POMODORO GRANULARE	Mar 17, 2023	Via dei Mercanti	Monza	10.50 €
MELISSA	Mar 17, 2023	Via dei Mercanti	Monza	17.50 €

Figura 3.23 Report parziale relativo al fatto Fornitura

```
SELECT p.Nome, t.Mese, o.Regione, f.Prezzounitario,  
SUM(f.Quantita) as Quantità  
FROM dim_prodotto p, dim_temporale t, dim_ordinecliente o, fact_vendita f  
WHERE p.id_prodotto = f.id_prodotto  
AND t.id_data=f.id_data  
AND o.id_ordinecliente = f.id_ordinecliente  
AND t.Anno= 2023  
AND o.Stato= "Italia"  
AND f.Prezzounitario<6  
GROUP BY p.Nome, t.Mese, o.Regione, f.Prezzounitario  
order by t.Mese;
```

Quantità di prodotti con prezzo unitario inferiore a 6 euro venduti in Italia nell'anno 2023:

Nome	Mese	Regione	Quantità	Prezzo unitario
Carciofi rustici al pepe nero	1	Toscana	21	€3.50
FOCU ML 212	1	Toscana	30	€5.50
La bruschetta alle olive e straccetti di acciughe ML106 BASSO	1	Toscana	27	€5.00
RUCOLA E BASTA KG3	1	Sicilia	52	€3.60
Sugo pronto all'Amatriciana 180 g.	1	Lazio	16	€5.50
TRIFOLATO FUNGHI PORCINI KG3	1	Puglia	9	€5.50
La bruschetta alle olive e straccetti di acciughe ML106 BASSO	1	Lombardia	12	€5.00
Crema desiderio	1	Trentino-Alto Adige	23	€5.50
Polpa pronta di pomodori 330g	1	Trentino-Alto Adige	30	€5.60
La bruschetta alle olive e straccetti di acciughe ML106 BASSO	1	Sicilia	26	€5.00
Vellutata di zucchine al profumo di mentuccia selvatica ml40	1	Sicilia	24	€3.50
Sugo pronto all'arrabbiata 330g	1	Puglia	6	€5.50
Trito di olive nere infornate e tocchetti di funghi porcini ML212 BASSO	1	Puglia	10	€3.90
Miele di millefiori ml40	1	Lazio	16	€5.50
Battuto di carciofi, prezzemolo e pepe bianco	1	Campania	12	€5.50
Ceci al rosmarino e cipolla	1	Campania	20	€5.50

Figura 3.24 Report parziale relativo al fatto Vendita

3.1.9 Dashboard

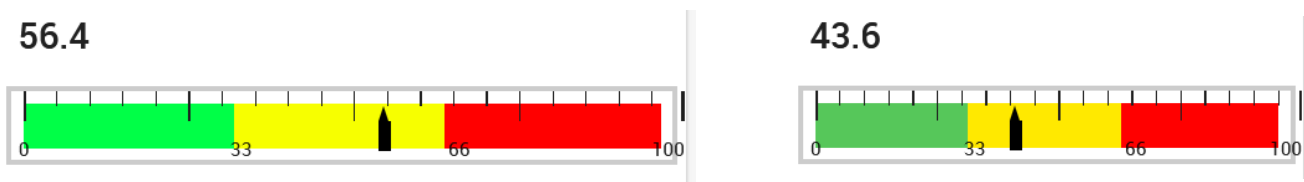
Le *dashboard* sono uno strumento di analisi che fornisce dati di sintesi tramite rappresentazioni grafiche di semplice comprensione.

Vengono rappresentati tre livelli di performance:

- basso, rappresentato dal colore rosso;
- accettabile, rappresentato dal colore giallo;
- buono, rappresentato dal colore verde.

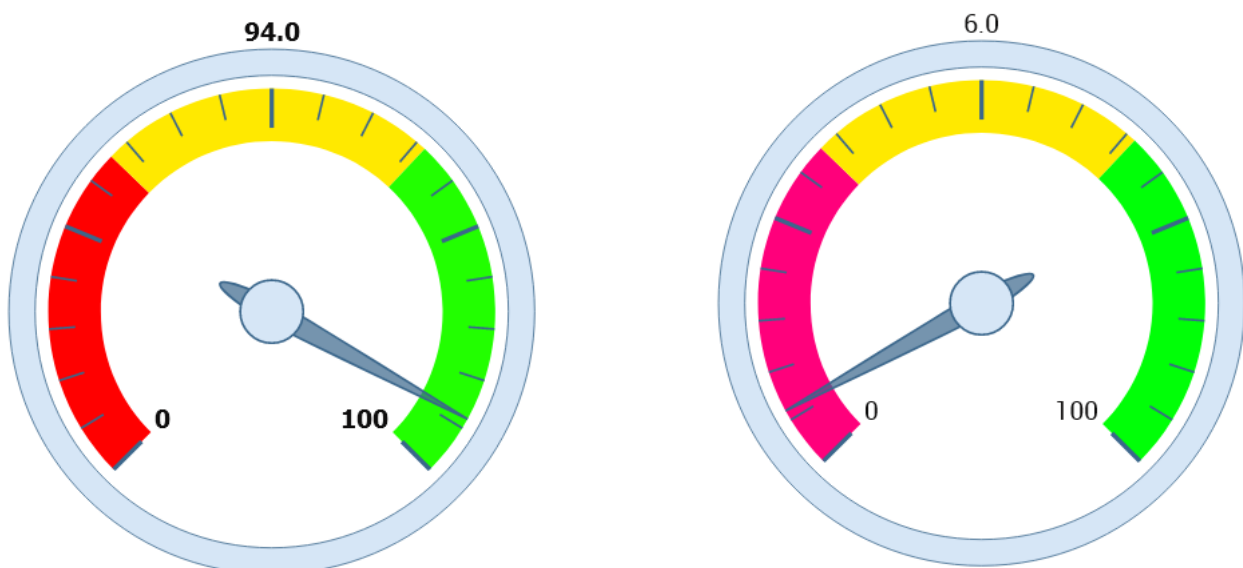
Di seguito, alcune dashboard di esempio.

Per determinare se l'azienda sta spendendo una quantità maggiore per i rifornimenti nel primo semestre rispetto al secondo semestre, è necessario confrontare le percentuali dell'importo speso per l'acquisto di materie prime nei due semestri:



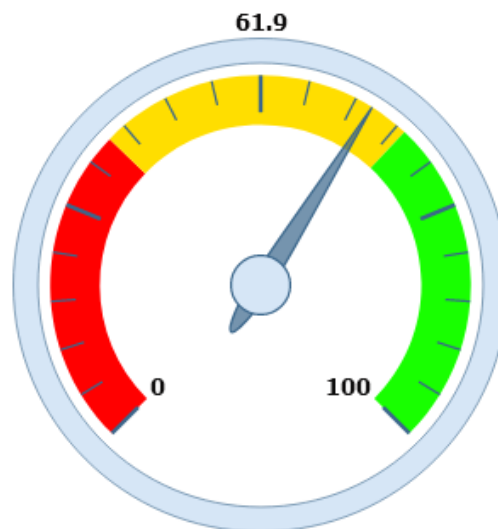
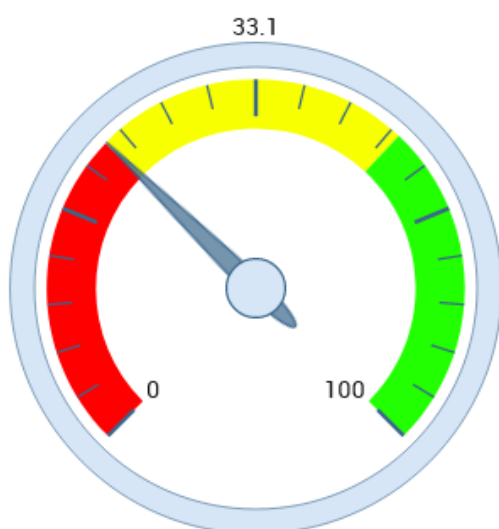
La situazione emerge come pressoché bilanciata, con una percentuale del 56,4% nel primo semestre e una percentuale del 43,6% nel secondo semestre.

Al fine di considerare l'opzione di centralizzare il processo di acquisto di materie prime in Italia, interrompendo le relazioni con i fornitori stranieri, si desidera confrontare rispettivamente la percentuale di materie prime acquistate in Italia con la percentuale di quelle acquistate all'estero.



Emergono chiaramente dati che indicano uno sbilanciamento a favore dell'Italia nel volume di materie prime acquistate. Pertanto, è più che plausibile considerare di trasferire integralmente il processo di rifornimento in Italia.

Al fine di valutare l'efficacia delle campagne pubblicitarie eseguite in Italia all'inizio dell'anno 2023, si intende confrontare la percentuale dell'importo ricavato in Italia nel 2022 con quella registrata nel 2023:



Certamente, ulteriori analisi sono necessarie per esaminare aspetti specifici, come ad esempio verificare l'assenza di una diminuzione dell'importo ricavato al di fuori dell'Italia. Tuttavia, il notevole aumento potrebbe verosimilmente essere attribuito all'efficacia delle campagne pubblicitarie.

Per concludere, l'obiettivo è visualizzare la percentuale di prodotti venduti durante il primo semestre nel Nord Italia e confrontarla con la percentuale di prodotti venduti, sempre nel Nord Italia, nel secondo semestre:

