

# OLAP con la suite Pentaho: *my first data warehouse system*



Luigi Bellomarini  
[luigi.bellomarini@uniroma3.it](mailto:luigi.bellomarini@uniroma3.it)



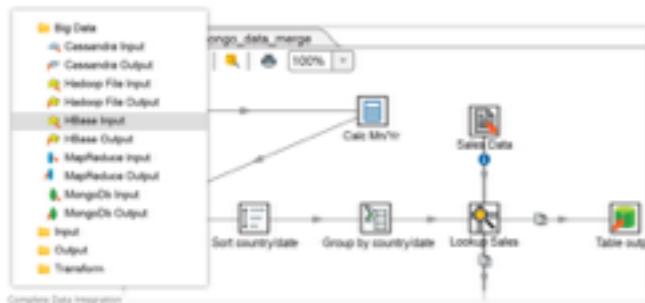
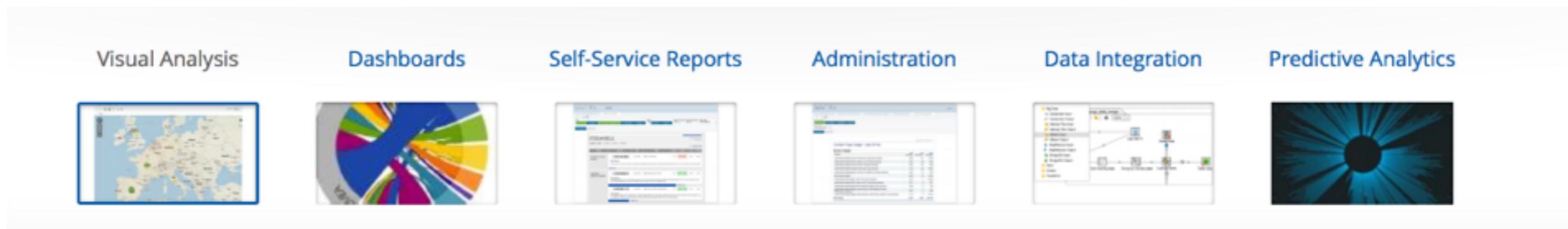
# Pro memoria

- Iscriversi al sito Moodle del corso!

# Sommario

- La suite Pentaho
  - Installazione e configurazione
  - Costruzione di un data mart
    - Interrogazioni OLAP
    - Cenni sull'architettura

# La suite



## ETL Data integration

## OLAP Business Analytics

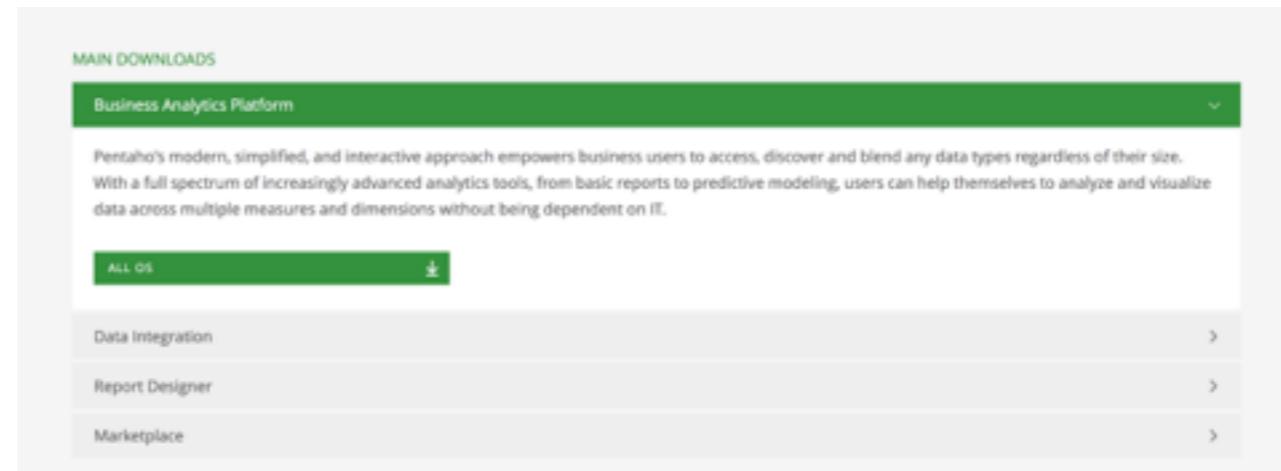


# Installazione e configurazione



# Installazione

- <http://community.pentaho.com>
  - **Pentaho Community 6.0**
    - Business Analytics
    - Data Integration
- Creare una directory “pentaho” (ad esempio /Applications/pentaho)
  - estrarre all’interno il contenuto dei due download
- **Requisiti di sistema**
  - JDK1.7.x (ad esempio 1.7.0\_45)
  - Chrome
  - 1.5 GB circa di spazio disponibile su disco



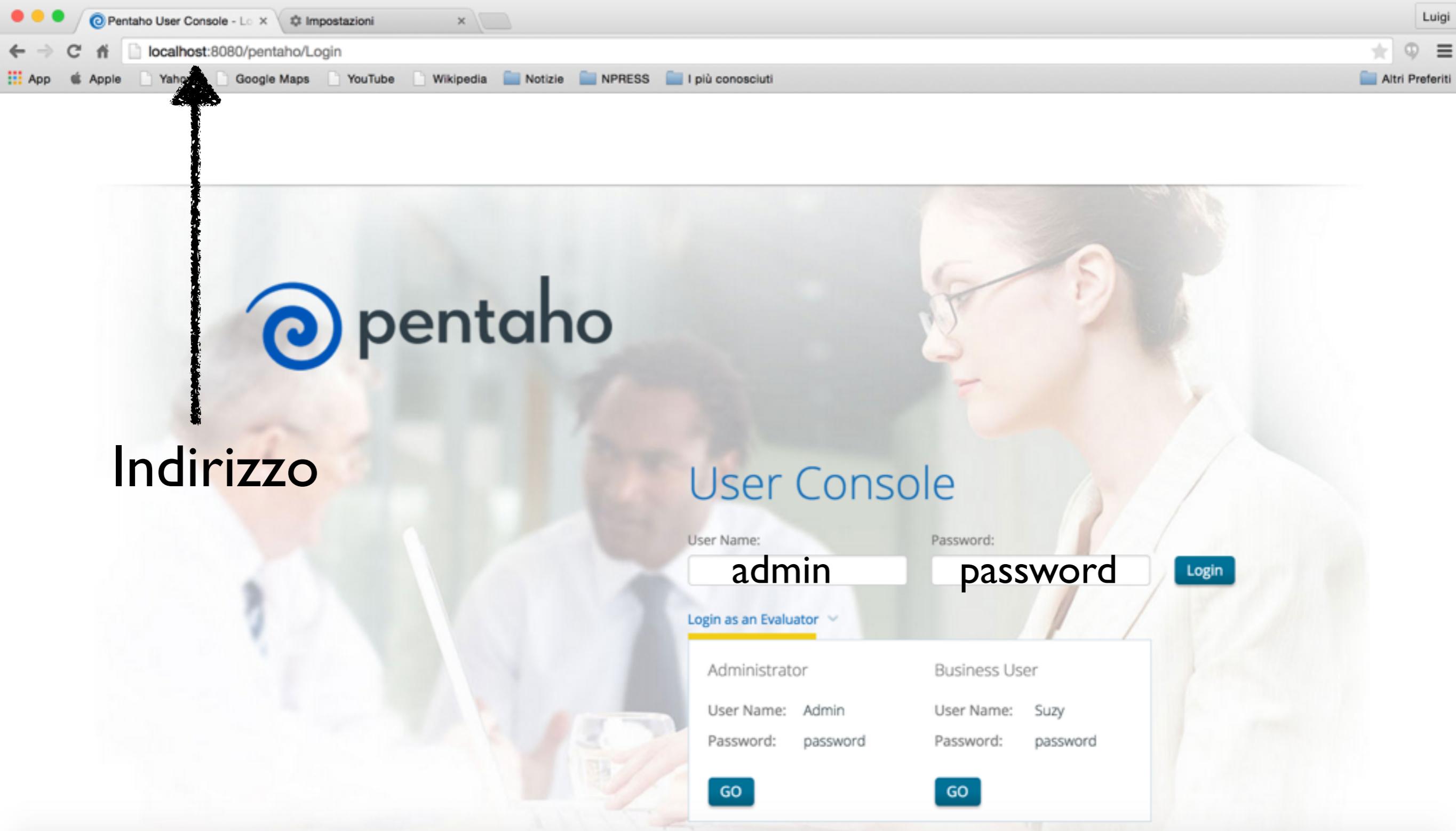
# Configurazione

- **PostgreSQL** nel PATH
  - ad esempio: `export PATH=$PATH:/Library/PostgreSQL/9.5/bin`
- **Driver JDBC** per Postgres (<https://jdbc.postgresql.org/download.html>)
  - Scegliere il driver più opportuno per la propria versione di Postgres
  - Copiare in `pentaho/biserver-ce-4/tomcat/lib`
- **export PENTAHO\_JAVA\_HOME=<path al proprio JDK>**
  - ad esempio: `export PENTAHO_JAVA_HOME=/Library/Java/JavaVirtualMachines/jdk1.7.0_45.jdk/Contents/Home`
- **Repository database (metadati)**
  - utilizzeremo il db HSQLDB già configurato in Business Analytics
  - non utilizzeremo Postgres per i metadati (configurazione laboriosa)

# Start/stop

- Posizionarsi nella directory pentaho/biserver-ce
  - `./start-pentaho.sh`
  - `./stop-pentaho.sh`
- I log sono in:
  - `.../tomcat/logs`
- Aprire Chrome
  - <http://localhost:8080>

# Business Analytics



# Business Analytics

File View Tools Help

Home admin

Browse Files

Create New

Manage Data Sources

Documentation

Recents

You haven't opened anything recently. Browse your files.

Browse Files

Favorites

You haven't

## Pentaho Business Analytics

Get help and contribute with your knowledge. Find here some ways how you can do it.

Live Chat

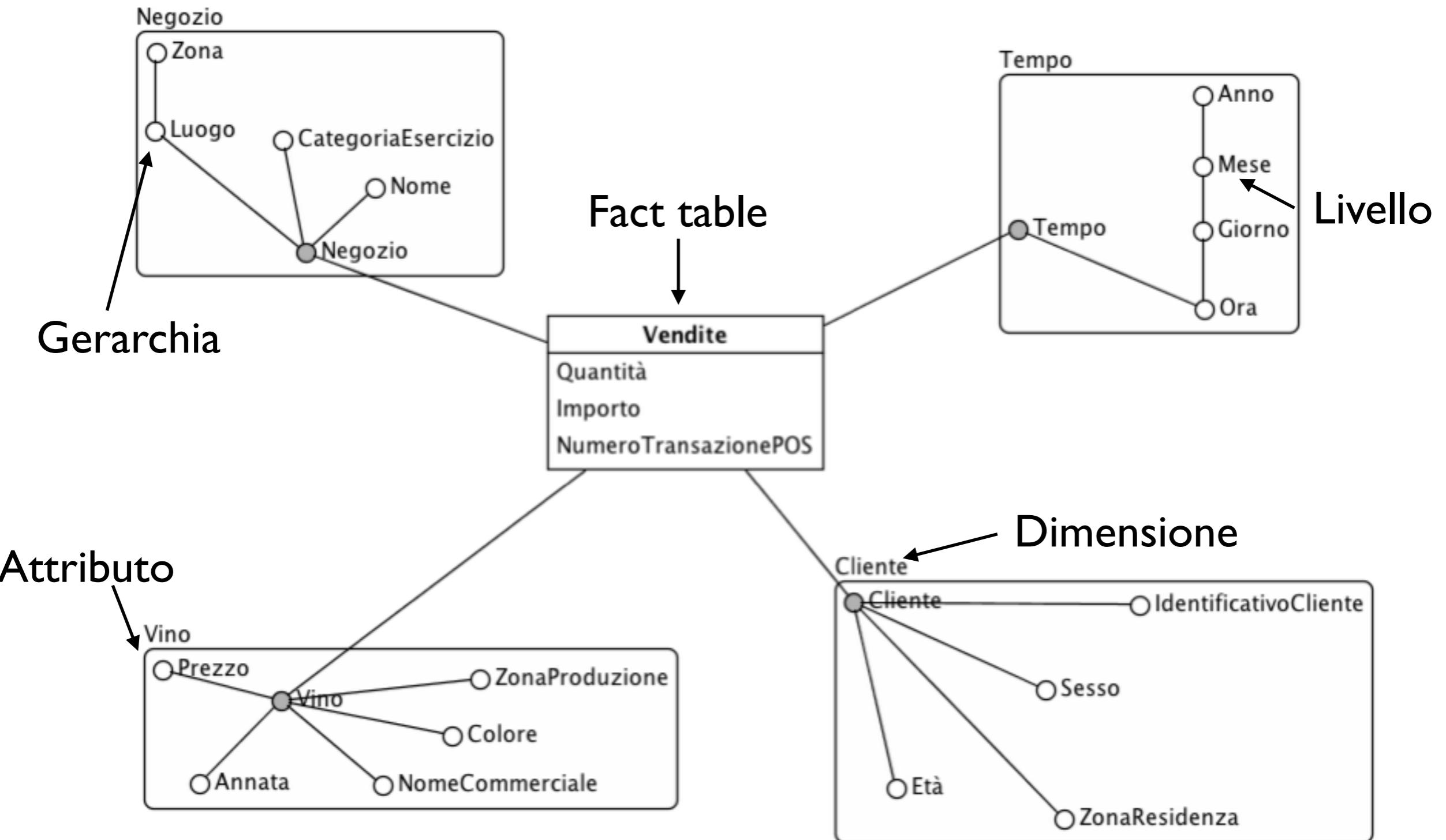


Documentation Forums Mailing Lists Blogs

# Costruzione e utilizzo di un data mart



# Schema dimensionale



# Attributo o livello gerarchico?

- Determinare se un nodo dello schema dimensionale è un attributo di un livello o un livello gerarchico.
  - Ragionare sulla **semantica** predecessore-successore
    - anno raggruppa i mesi, mentre Nome non i negozi.
  - Ragionare sul **contenimento** tra predecessore e successore
    - un anno è un insieme di mesi
    - Il nome di un negozio non è un insieme di negozi
  - Ragionare sulla **molteplicità** dell'associazione predecessore-successore
    - N mesi, I anno (**livello**)
    - I nome negozio, I negozio (**attributo**).

# Schema dimensionale

- **Vendite(k\_negozi, k\_tempo, k\_cliente, k\_vino, quantita, importo, numero\_transazione\_pos)**
- **Negozi(k\_negozi, nome, luogo, categoria\_esercizio)**
- **Tempo(k\_tempo, anno, mese, giorno, ora, ...)**
- **Cliente(k\_cliente, eta, zona\_residenza, sesso)**
- **Vino(k\_vino, prezzo, nome\_commerciale, colore, zona\_produzione, annata)**

# Obiettivi

- Saper costruire nello strumento Pentaho un data mart
- Eseguire alcune semplici interrogazioni
- Costruire interfacce per analizzare i dati

# Creazione schema dimensionale

- Creare in Postgres un database con nome **DWH**
- Creare in DWH lo schema **Wines**
- Creare in Wines le varie relazioni

# Creazione db

The screenshot shows a database management interface with two main panes. The left pane is a tree view of database objects:

- FTS Dictionaries (0)
- FTS Parsers (0)
- FTS Templates (0)
- Functions (0)
- Sequences (4): cliente\_k\_cliente\_seq, negozio\_k\_negozio\_seq, tempo\_k\_tempo\_seq, vino\_k\_vino\_seq
- Tables (5):
  - cliente:
    - Columns (5): eta, zona\_residenza, sesso, identificativo, k\_cliente
    - Constraints (1)
    - Indexes (0)
    - Rules (0)
    - Triggers (0)
  - negozio:
    - Columns (5): luogo, zona, categoria\_esercizio, nome, k\_negozio
    - Constraints (1)
    - Indexes (0)
    - Rules (0)
    - Triggers (0)

The right pane displays the properties of the 'vendite' table and its SQL definition:

Property	Value
Name	vendite
OID	16632
Owner	postgres
Tablespace	pg_default
ACL	
Of type	
Primary key	k_tempo, k_vino, k_cliente, k_negozio
Rows (estimated)	0
Fill factor	
Rows (counted)	0

SQL pane:

```
-- Table: "Wines".vendite
-- DROP TABLE "Wines".vendite;

CREATE TABLE "Wines".vendite
(
    quantita numeric,
    importo numeric,
    numero_transazione_pos numeric,
    k_tempo integer NOT NULL,
    k_vino integer NOT NULL,
    k_cliente integer NOT NULL,
    k_negozio integer NOT NULL,
    CONSTRAINT vendite_pkey PRIMARY KEY (k_tempo, k_vino, k_cliente, k_negozio),
    CONSTRAINT vendite_k_cliente_fkey FOREIGN KEY (k_cliente)
        REFERENCES "Wines".cliente (k_cliente) MATCH SIMPLE
        ON UPDATE NO ACTION ON DELETE NO ACTION,
    CONSTRAINT vendite_k_negozio_fkey FOREIGN KEY (k_negozio)
        REFERENCES "Wines".negozi (k_negozio) MATCH SIMPLE
        ON UPDATE NO ACTION ON DELETE NO ACTION,
    CONSTRAINT vendite_k_tempo_fkey FOREIGN KEY (k_tempo)
        REFERENCES "Wines".tempo (k_tempo) MATCH SIMPLE
        ON UPDATE NO ACTION ON DELETE NO ACTION,
    CONSTRAINT vendite_k_vino_fkey FOREIGN KEY (k_vino)
        REFERENCES "Wines".vino (k_vino) MATCH SIMPLE
)
```

# Creazione DB (convenzioni)

- Nomi minuscoli per relazioni, colonne, etc.
  - `set search_path to "Wines"`
  - si possono eseguire interrogazioni senza quotes
- **Tipi dato**
  - `text`
  - `numeric`
  - `serial` (dimensioni) vs `integer` (fact table)
- Sequence

# Dati

- Tralasciamo per il momento il popolamento del database
- Tratteremo l'ETL successivamente

The screenshot shows the pgAdmin III interface. On the left, there are two tabs: "clienti.csv" and "vini.txt". The "vini.txt" tab contains the following CSV data:

	zona_produzione	nome_commerciale	colore	annata	prezzo
1	la segreta	dispensa	bianco	2002	7
2	la segreta	dispensa	bianco	2003	8
3	la segreta	dispensa	bianco	2004	5
4	la segreta	dispensa	bianco	2005	6
5	la segreta	dispensa	bianco	2006	7
6	rosé	dispensa	rosa	2010	7
7	rosé	dispensa	rosa	2011	6
8	rosé	dispensa	rosa	2012	7
9	rosé	dispensa	rosa	2013	6
10	rosé	dispensa	rosa	2014	6
11	burdese	dispensa	rosso	2014	18
12	cometa	dispensa	giallo	2013	18
13	cometa	dispensa	giallo	2012	20
14	cometa	dispensa	giallo	2011	19
15	cometa	dispensa	giallo	2010	17
16	frappato	dorilli	rosso	2009	9
17	frappato	dorilli	rosso	2010	10
18	frappato	dorilli	rosso	2011	11
19	frappato	dorilli	rosso	2012	7
20	cerasuolo di vittoria	dorilli	rosso	2012	11
21	cerasuolo di vittoria	dorilli	rosso	2013	12
22	cerasuolo di vittoria	dorilli	rosso	2014	11

The right side of the interface shows the pgAdmin Object browser for the "vino" table. The browser displays the following properties for the "vino" table:

Property	Value
Name	vino
OID	16608
Owner	postgres

The "Properties" tab also shows the table's constraints, indexes, rules, and triggers.

The "Data" tab displays the actual data from the "vino" table, which matches the data shown in the "vini.txt" tab. The table has columns: prezzo (numeric), nome\_commerciale (text), colore (text), zona\_produzione (text), k\_vino (integer PK), and annata (numeric).

# Creazione Data Source

File View Tools Help

Home admin

Browse Files

Create New

Manage Data Sources

Documentation

Recents

You haven't opened anything recently. Browse your files.

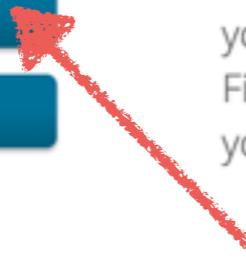
Browse Files

Favorites

You haven't

## Pentaho Business Analytics

Get help and contribute with your knowledge. Find here some ways how you can do it.



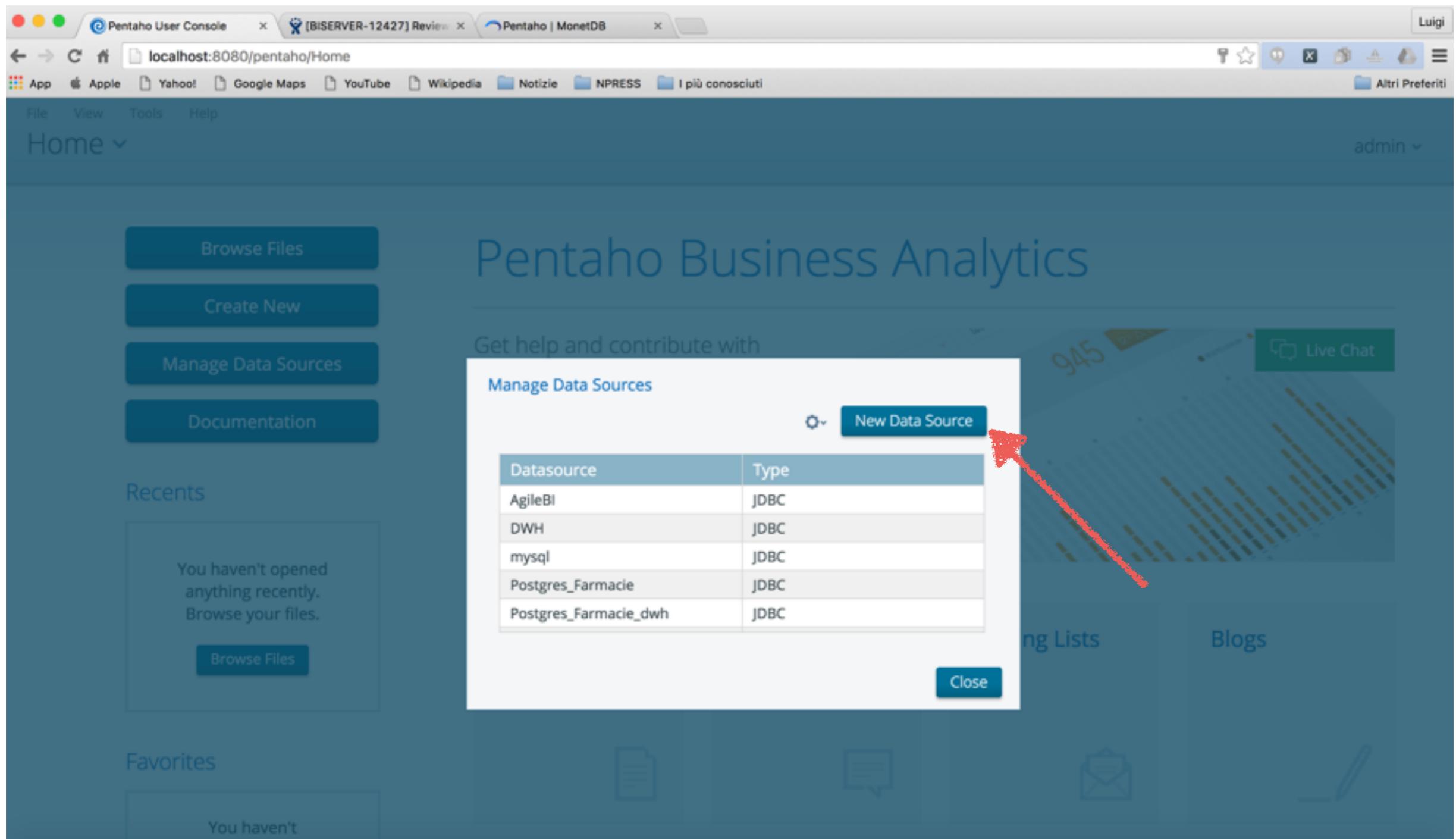
Documentation

Forums

Mailing Lists

Blogs

# Creazione Data Source



# Creazione Data Source

Sales View Tools Help

Home admin

Browse Create Manage Data Document Live Chat

Recents You haven't anything recently. Browse your recent files.

Favorites You have no favorites yet. Browse your favorite files.

Data Source Wizard

Select Source Type

Select Tables

Define Joins

Data Source Name: Wines

Source Type: Database Table(s)

Select a database connection and click Next to choose from a list of the available database tables.

Connection:

AgileBI (highlighted with a red arrow)

DWH

Postgres\_Farmacie

Postgres\_Farmacie\_dwh

SampleData

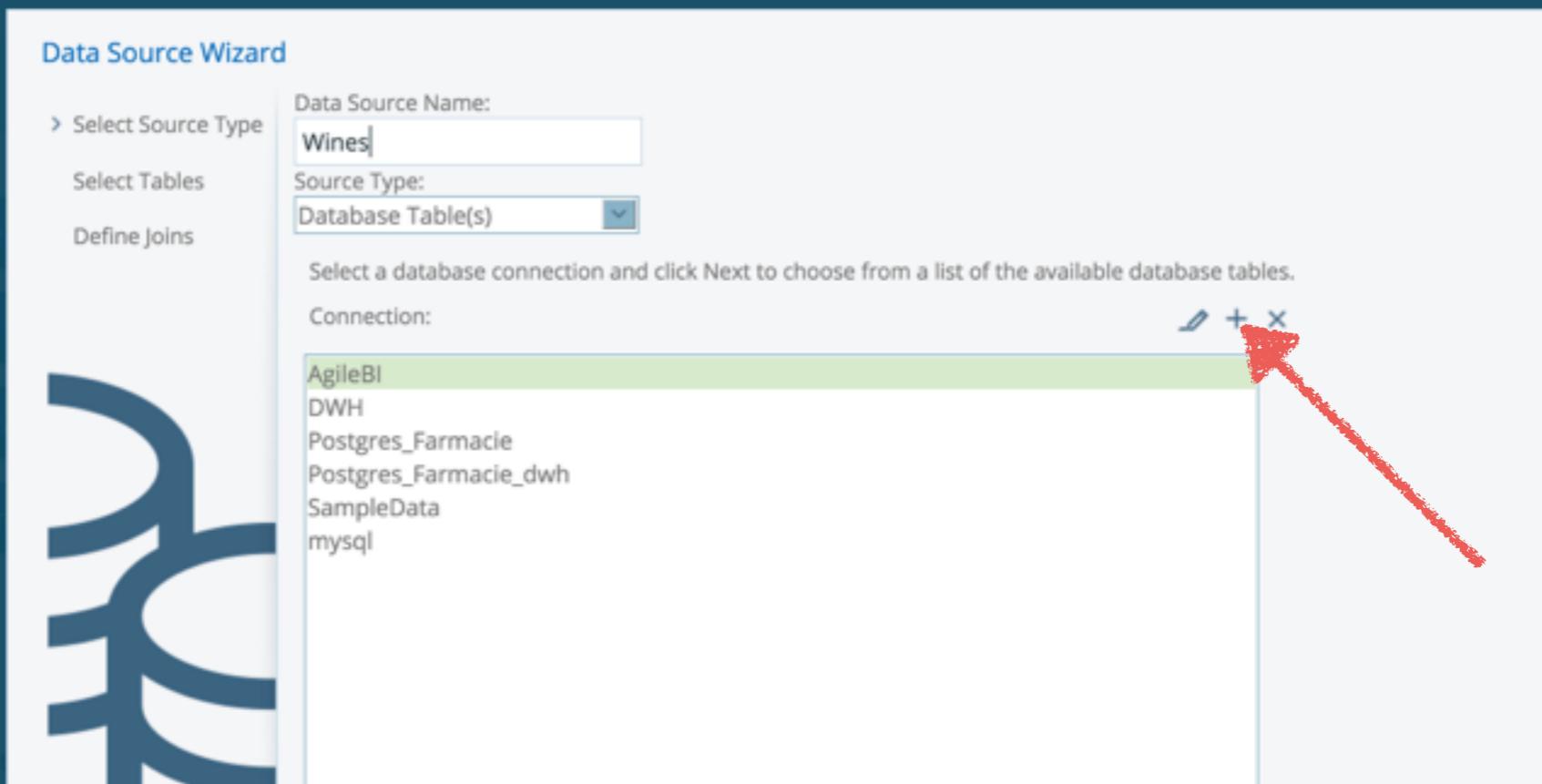
mysql

Create data source for:

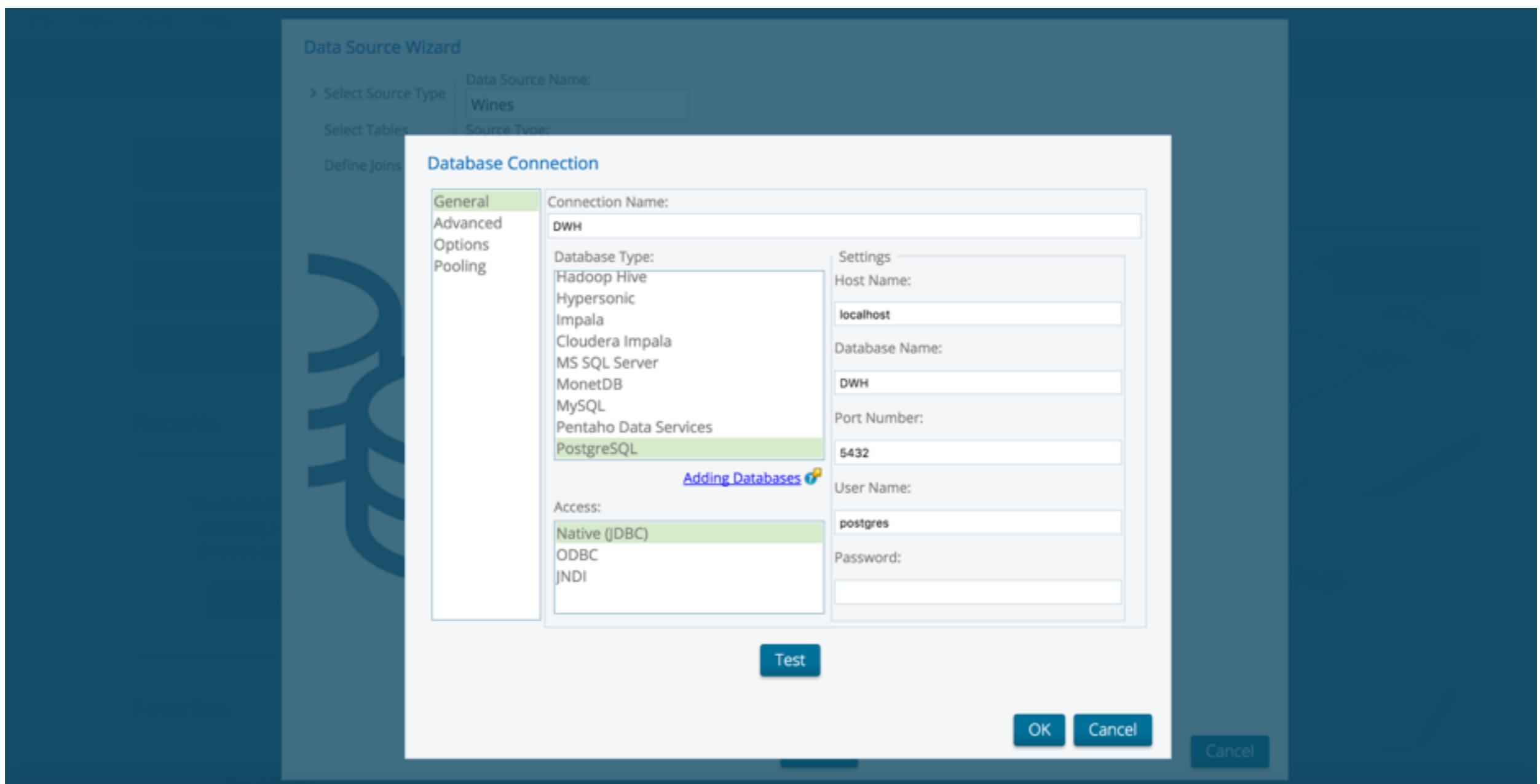
Reporting only

Reporting and Analysis (Requires Star Schema)

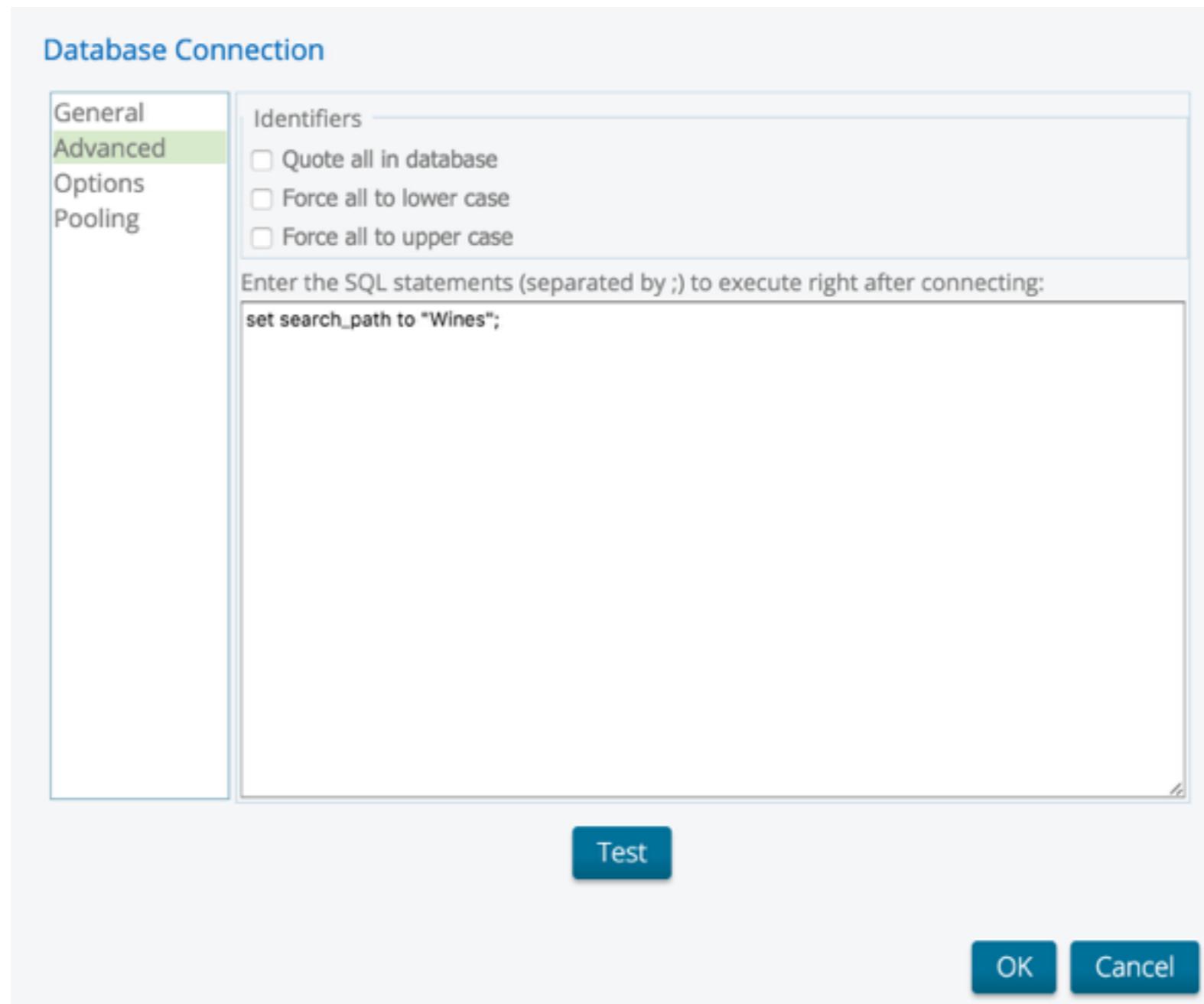
< Back Next > Finish Cancel



# Creazione Data Source



# Creazione Data Source



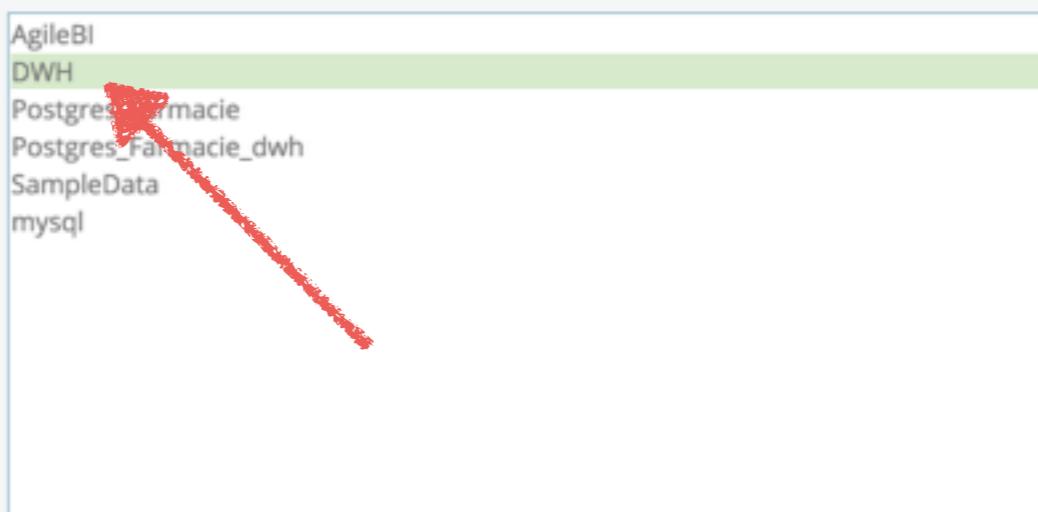
# Creazione Data Source

**Data Source Wizard**

» Select Source Type  
Select Tables  
Define Joins

Data Source Name:   
Source Type:

Select a database connection and click Next to choose from a list of the available database tables.

Connection:   
AgileBI  
DWH  
Postgres\_Farmacie  
Postgres\_Farmacie\_dwh  
SampleData  
mysql

Create data source for:  
 Reporting only  
 Reporting and Analysis (Requires Star Schema)

< Back  Finish



# Fact table

Home > Data Sources > New

admin -

**Data Source Wizard**

Select Source Type

> Select Tables

Define Joins

Browse

Create

Manage Data

Document

Recents

You haven't created anything yet. Browse your organization's data.

Browse

Favorites

My Profile

Live Chat

Blogs

Schema: Wines

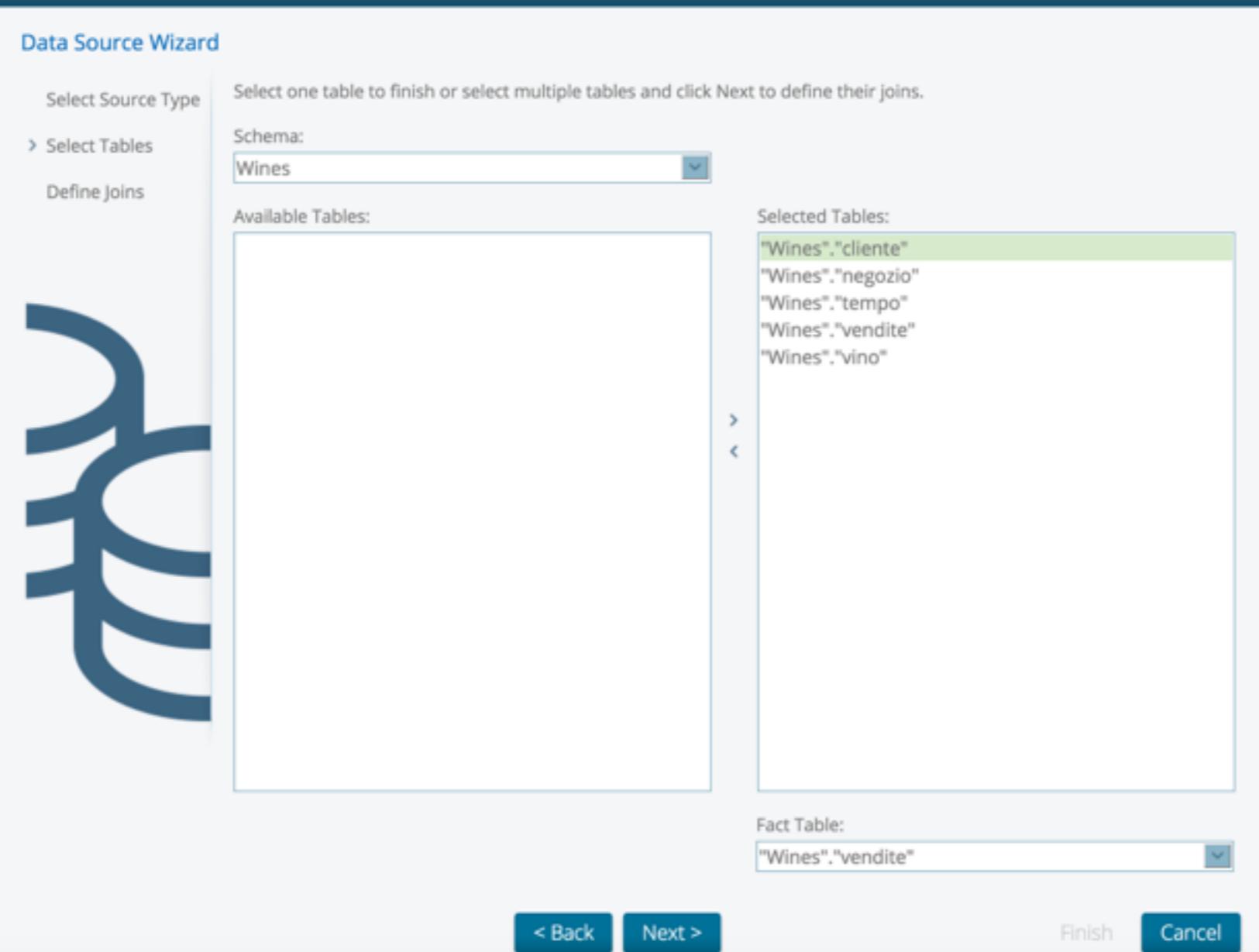
Available Tables:

Selected Tables:

- "Wines"."cliente"
- "Wines"."negozi"
- "Wines"."tempo"
- "Wines"."vendite"
- "Wines"."vino"

Fact Table: "Wines"."vendite"

< Back Next > Finish Cancel



# Percorsi di join

**Data Source Wizard**

Select Source Type      Define how the tables join to each other. All tables must have at least one join defined.

Select Tables      > Define Joins

Left Table: "Wines"."vendite"      Right Table: "Wines"."negozi"

Key Field: quantita, importo, numero\_transazione\_pos, k\_tempo, k\_vino, k\_cliente, k\_negozi

Key Field: luogo, zona, categoria\_esercizio, nome, k\_negozi

[Create Join](#)

Join(s): "Wines"."vendite".k\_cliente - INNER JOIN - "Wines"."cliente".k\_cliente

[Delete Join](#)

< Back      Next >      [Finish](#)      [Cancel](#)

# Definizione schema dimensionale

Data Source Model Editor

The screenshot shows the 'Data Source Model Editor' interface. On the left, the 'Available' pane lists various data items under categories like Cliente, Negozio, and Tempo. In the center, the 'Analysis' tab is active, displaying a hierarchical structure of 'Measures' and 'Dimensions'. The 'Dimensions' section contains four entries: Cliente, Negozio, Tempo, and Vino, each preceded by a blue double-headed arrow icon. The right side of the screen is a 'Properties' panel which is currently empty. At the bottom right are 'OK' and 'Cancel' buttons.

Available

- Cliente
  - Eta
  - Identificativo
  - K cliente
  - Sesso
  - Zona residenza
- Negozio
  - Categoria esercizio
  - K negozio
  - Luogo
  - Nome
  - Zona
- Tempo
  - Anno
  - Giorno

Analysis Reporting

- Measures
- Dimensions
  - Cliente
  - Negozio
  - Tempo
  - Vino

Properties

OK Cancel

# Definizione schema dimensionale

Data Source Model Editor

Available

- Cliente
  - Eta
  - Identificativo
  - K cliente
  - Sesso
  - Zona residenza
- Negozio
  - Categoria esercizio
  - K negozio
  - Luogo
  - Nome
  - Zona
- Tempo
  - Anno
  - Giorno

Analysis Reporting

Dimensions

Cliente

Properties

Source Column: Identificativo

Details: Level Name: Cliente

Contains only unique members

Ordinal Column:

Geography Type: None

Gerarchia

Dimensione

Livello

OK Cancel

# Mapping schema dimensionale - schema relazionale

Data Source Model Editor

The screenshot shows the Data Source Model Editor interface. On the left, the 'Available' pane lists dimensions: Cliente, Negozio, and Tempo. The 'Analysis' pane shows the dimensional schema, and the 'Reporting' pane shows the relational schema. Arrows point from the 'Available' dimensions to their corresponding objects in the schema. The 'Properties' pane on the right contains settings for the selected object.

**Available**

- Cliente
  - Eta
  - Identificativo
  - K cliente
  - Sesso
  - Zona residenza
- Negozio
  - Categoria esercizio
  - K negozio
  - Luogo
  - Nome
  - Zona
- Tempo
  - Anno
  - Giorno

**Analysis**

- Measures
- Dimensions
  - Cliente
    - Cliente
    - Eta
    - Sesso
    - Zona residenza
  - Negozio
    - Negozio
    - Categoria esercizio
    - Luogo
    - Zona
  - Tempo
    - Tempo
    - Tempo

**Reporting**

- Measures
- Dimensions
  - Cliente
  - Eta
  - Sesso
  - Zona residenza
- Negozio
- Tempo

**Properties**

Source Column: Identificativo

Details: Level Name: Cliente

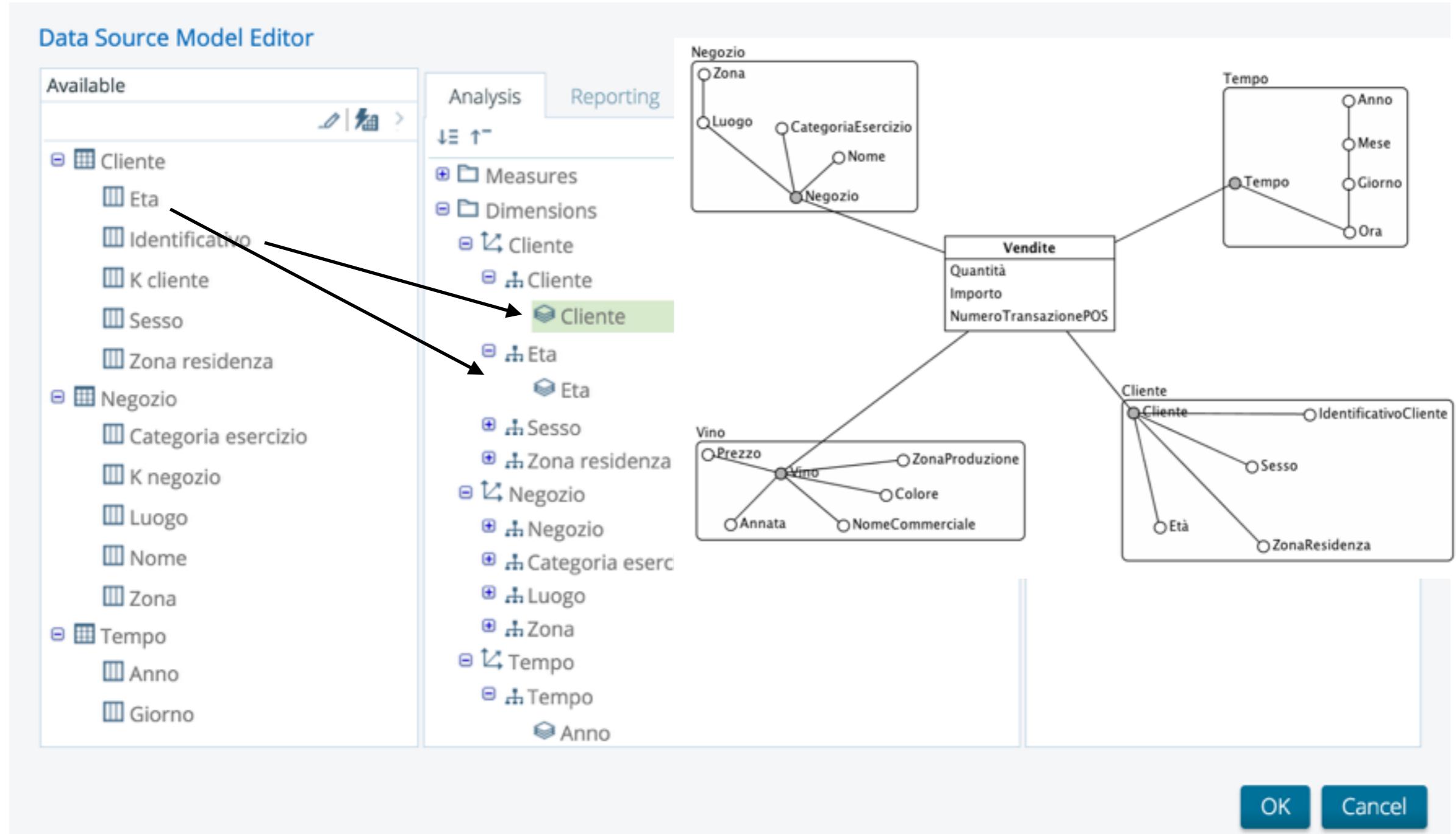
Contains only unique members

Ordinal Column:

Geography Type: None

OK Cancel

# Mapping schema dimensionale - schema relazionale



# Data

Data Source Model Editor

**Data**

```
graph LR; Anno --- Mese; Mese --- Giorno; Giorno --- Data
```

Zona  
Tempo  
Anno  
Giorno  
K tempo  
Mese  
Vendite  
Importo  
K cliente  
K negozio

**Reporting**

- ↓ Zona residenza
- ↓ Negozio
- ↓ Categoria esercizio
- ↓ Luogo
- ↓ Zona
- ⊖ ↗ Tempo
- ⊖ ↗ Vino
- ⊕ ↗ Vino
- ⊕ ↗ Annata
- ⊕ ↗ Colore
- ⊕ ↗ Prezzo
- ⊕ ↗ Zona produzione

**Properties**

Source Column: Anno

Details: Level Name: Anno

Contains only unique members

Ordinal Column:

Time Level Type: Years

Source Column Format: yyyy

OK Cancel

# Indicazioni

- Si creano **misure e dimensioni**
- Per ogni dimensione si creano delle **gerarchie**
- Si crea una **gerarchia di default** con lo stesso nome della dimensione
- Per la gerarchia di default, si crea un primo **livello di default** con lo stesso nome, legato ad un attributo relazionale sufficientemente descrittivo, possibilmente univoco e sintetico (non alla chiave)
- Si possono creare altre gerarchie, con **vari livelli**
- Gli attributi hanno un **livello di appartenenza**
- Conveniente utilizzare gerarchie e livelli per tutti gli attributi relazionali sui quali si intende fare delle interrogazioni (potatura e innesto)
- In sostanza, conviene **usare gli attributi in modo limitato**

# Interrogazioni

Screenshot of the Pentaho User Console - JF window showing the Pentaho Business Analytics homepage.

The browser title bar reads "Pentaho User Console - JF". The address bar shows "localhost:8080/pentaho/Home". The user is logged in as "Luigi" under "admin".

The main content area features the "Pentaho Business Analytics" logo and several navigation links:

- Browse Files
- Create New
- Manage Data Sources
- Documentation
- CDE Dashboard
- Jpivot View
- Data Source

A "Try Enterprise Edition" button is located in the top right corner.

The "Data Source" link is currently selected, highlighted with a yellow bar. A tooltip provides information about the Pentaho Business Analytics Platform:

Pentaho Business Analytics Platform delivers an integrated, end-to-end solution including data integration, visualization and consumption of data. Either by providing ad-hoc tools to analyze and visualize data or by serving pre-created content, different content types including analysis, reports, dashboards, data mining and even community created plugins can be seamlessly connected.

To the right of the text is a large image of a computer monitor displaying a dashboard interface.

On the left side, there are sections for "Recents" and "Favorites". The "Recents" section lists a recent item: "denormalized-mo...". The "Favorites" section indicates "You haven't added any favorites yet".

At the bottom, there are links for "Getting Started", "Documentation", and "Forums".



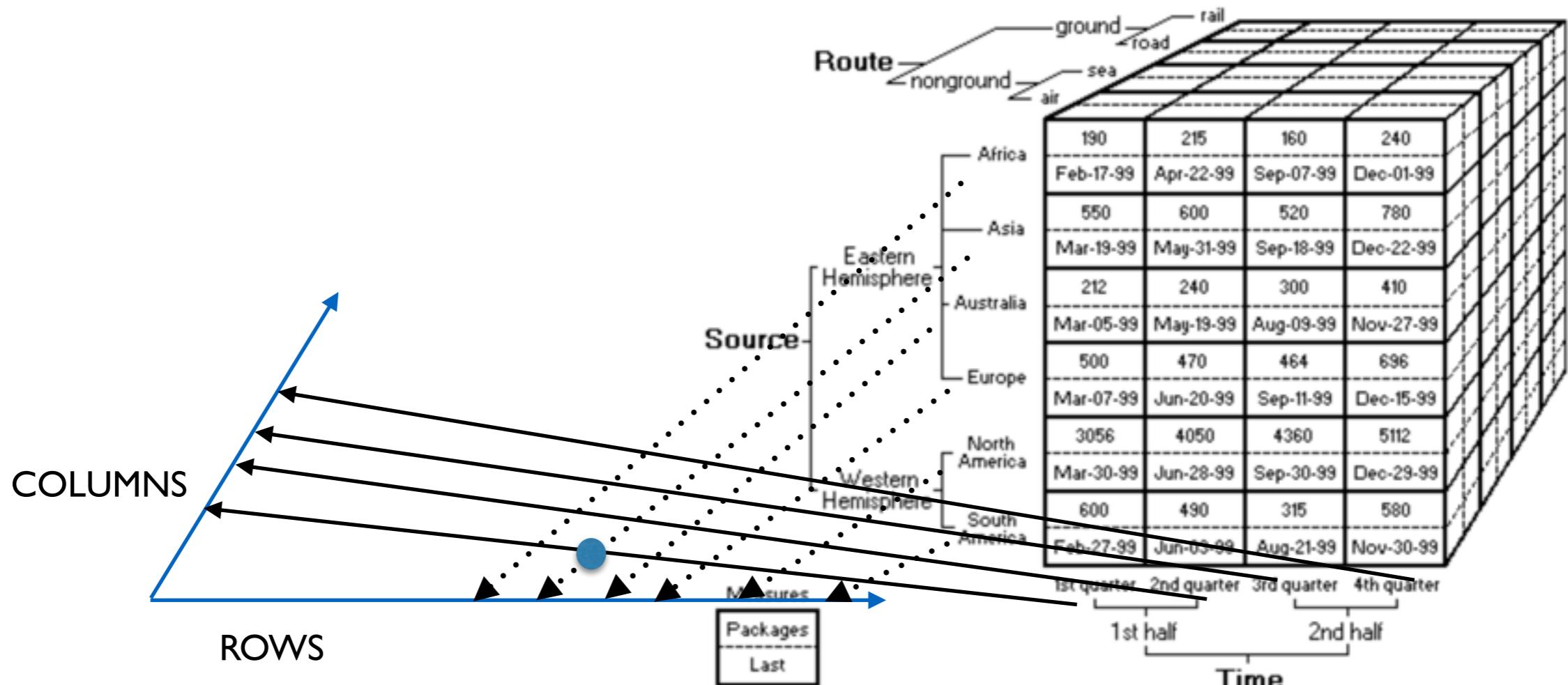
# MDX

- MultiDimensionalExpressions.
- Linguaggio di interrogazione per DW dimensionali.
- Le espressione MDX si possono tradurre in query SQL (laboriose).
- Ideato dal 1997 da Microsoft nell'ambito del progetto di SQLServer.
- Adottato dalla maggior parte dei sistemi OLAP.
- Standard de facto per i DW dimensionali.

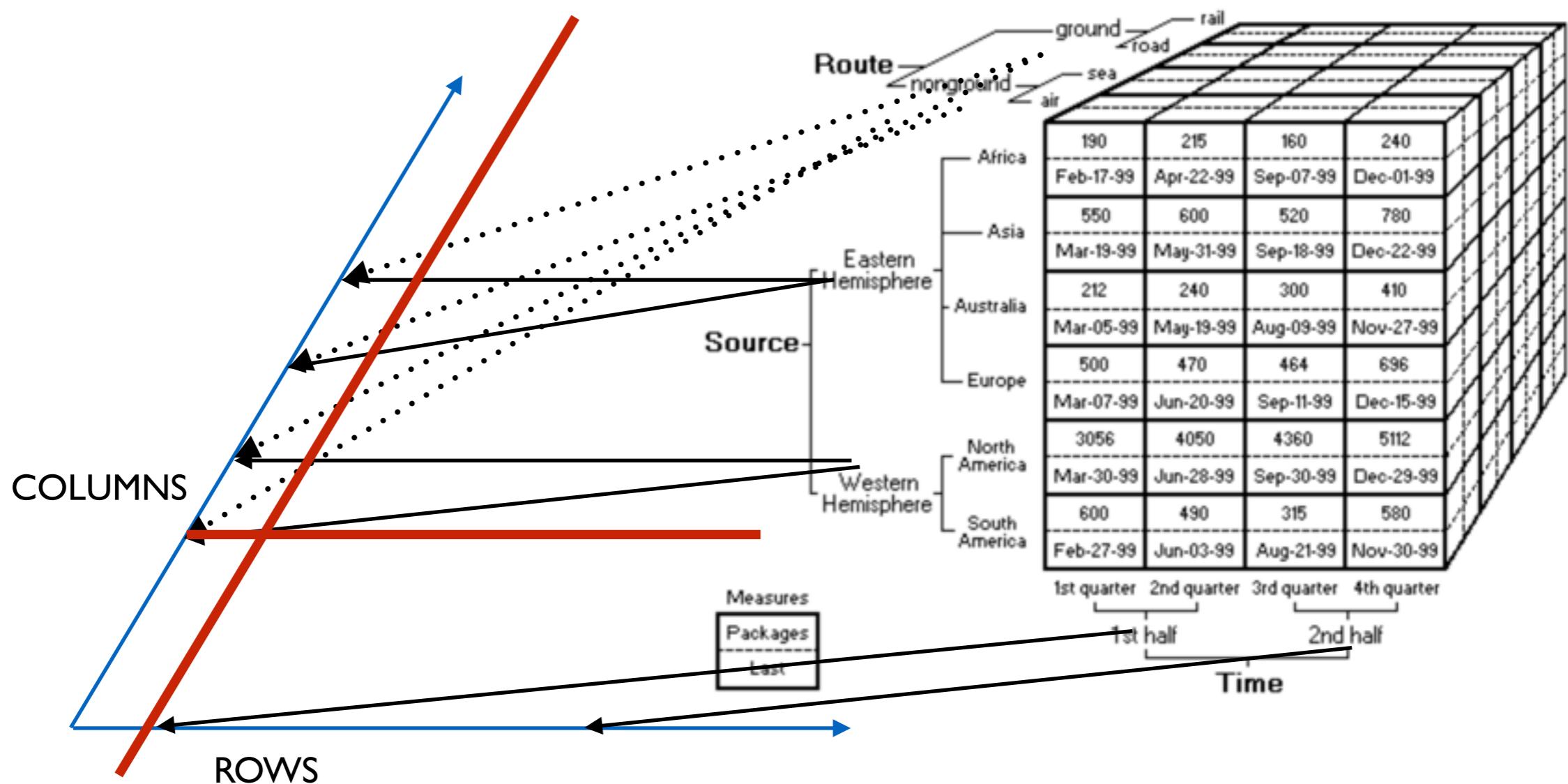
# MDX

- Sintassi in apparenza simile a SQL, attenzione.
  - Clauses **SELECT**, **FROM**, **WHERE** con significato diverso.
- Mentre SQL produce un result set di struttura tabellare (non necessariamente una relazione), MDX produce un **cubo multidimensionale**.
- Risulta conveniente immaginare le dimensioni del cubo di output come assi.
  - Due o più assi
  - Una query MDX mappa gli elementi delle dimensioni sugli assi di un cubo multidimensionale di output.
  - Per ogni combinazione dei punti degli assi, calcola l'aggregato delle misure.
  - Scrivere una query MDX equivale a **definire i punti degli assi del cubo di output**.
  - Mappare N dimensioni su M assi.

# Mappatura di due dimensioni in due assi



# Mappatura di tre dimensioni in due assi

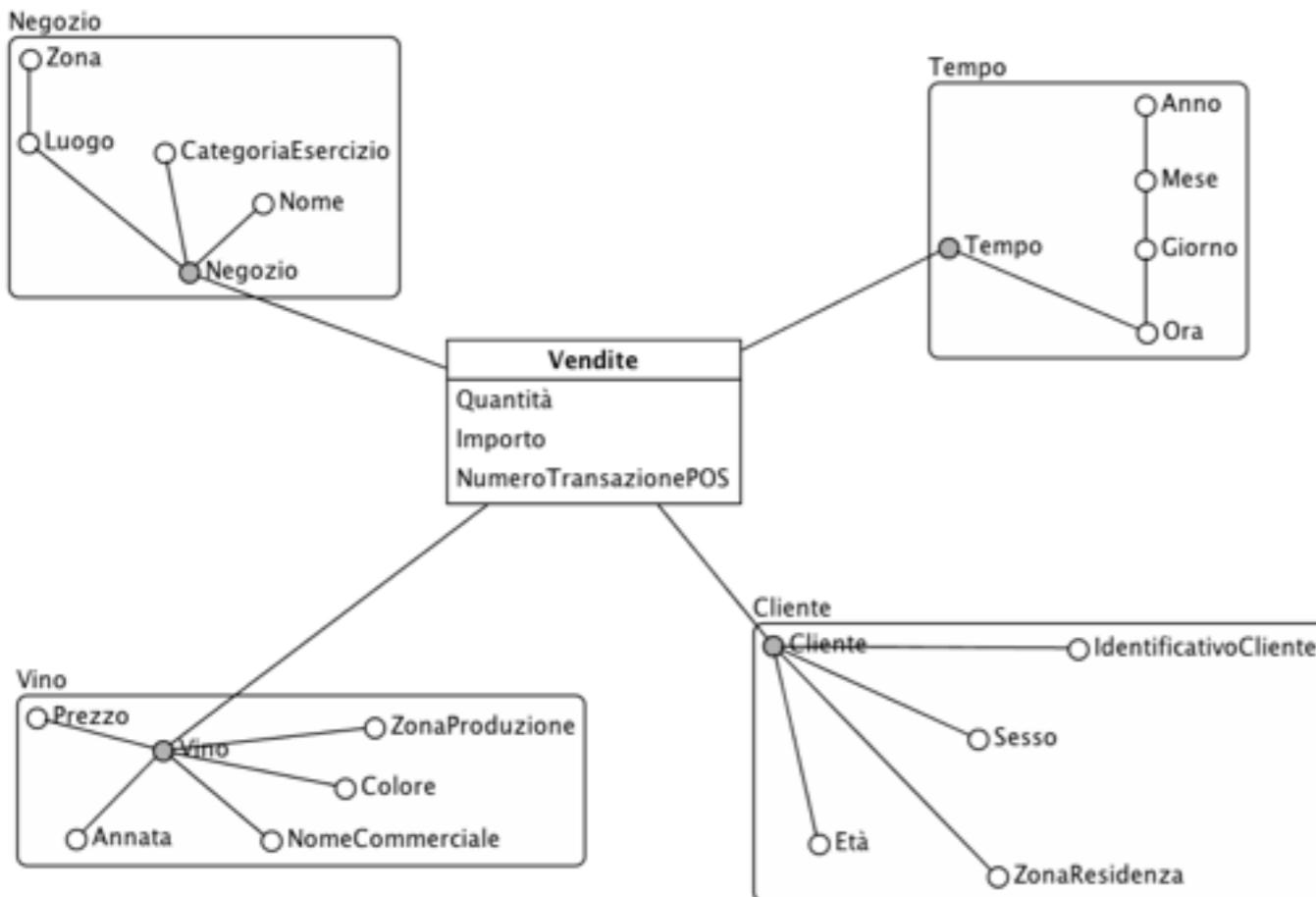


# Meta-definizioni e definizioni MDX

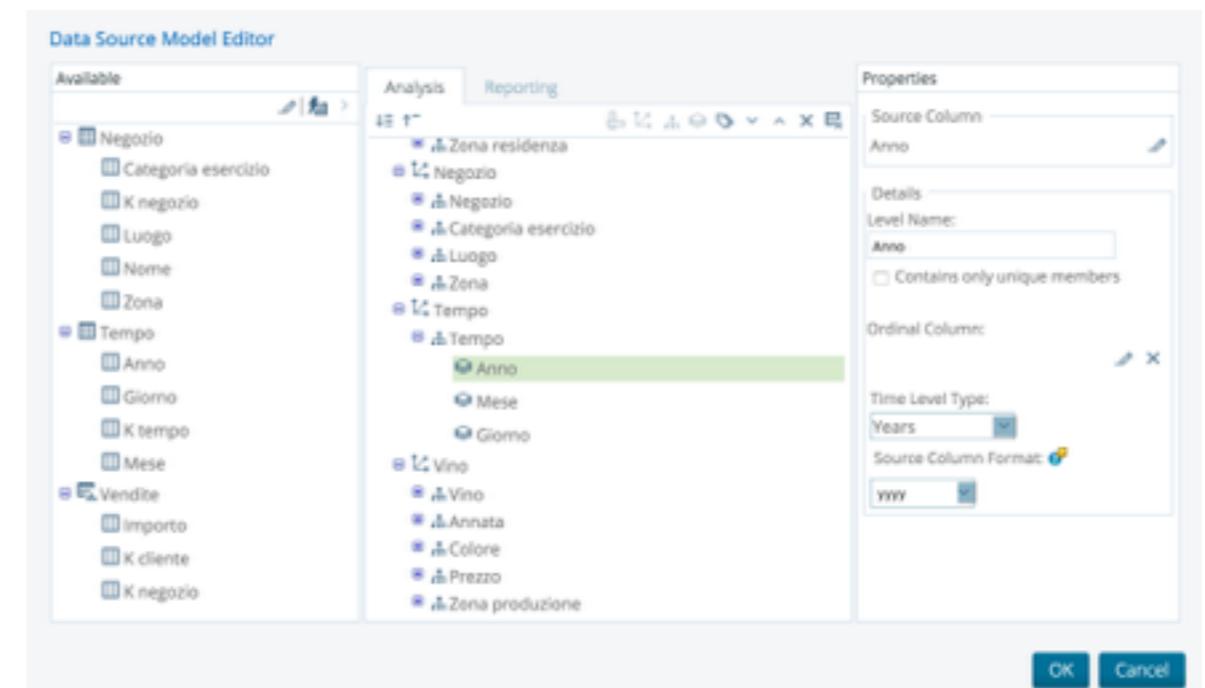
- **Meta-definizioni:**
  - come referenziare dimensioni, gerarchie e livelli.
- **Definizioni:**
  - come referenziare elementi veri e propri delle dimensioni (membri)
  - come accedere ai loro attributi.

# Meta-definizioni

## Nome di dimensione

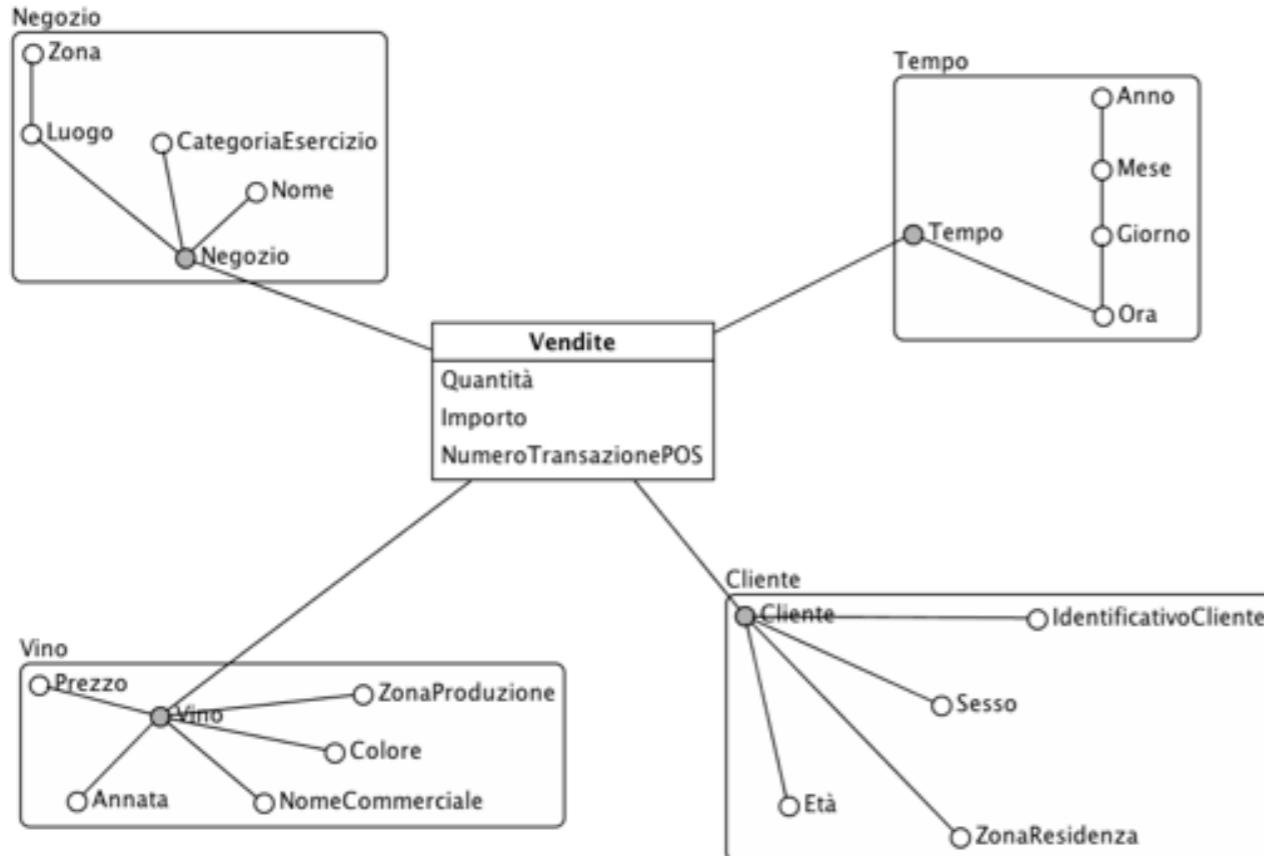


- Scelta di una dimensione di analisi.
  - [Dimensione]
- Esempi
  - [Tempo]
  - [Negozio]
  - [Vino]

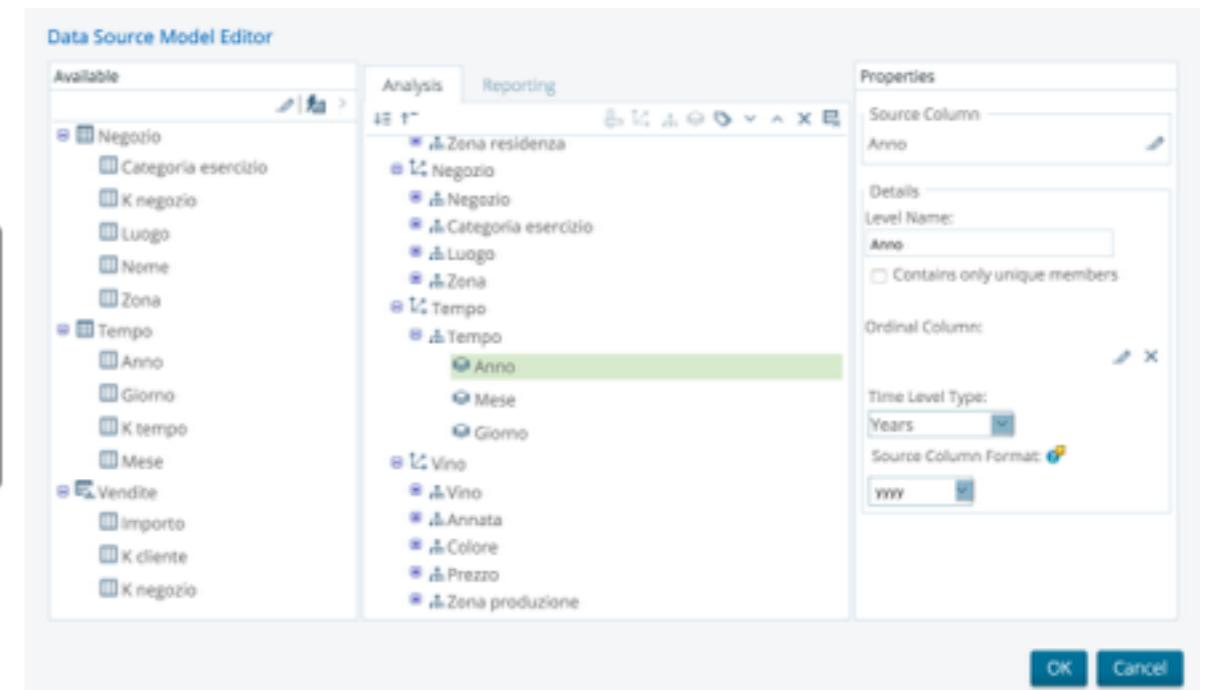


# Meta-definizioni

## Nome di gerarchia

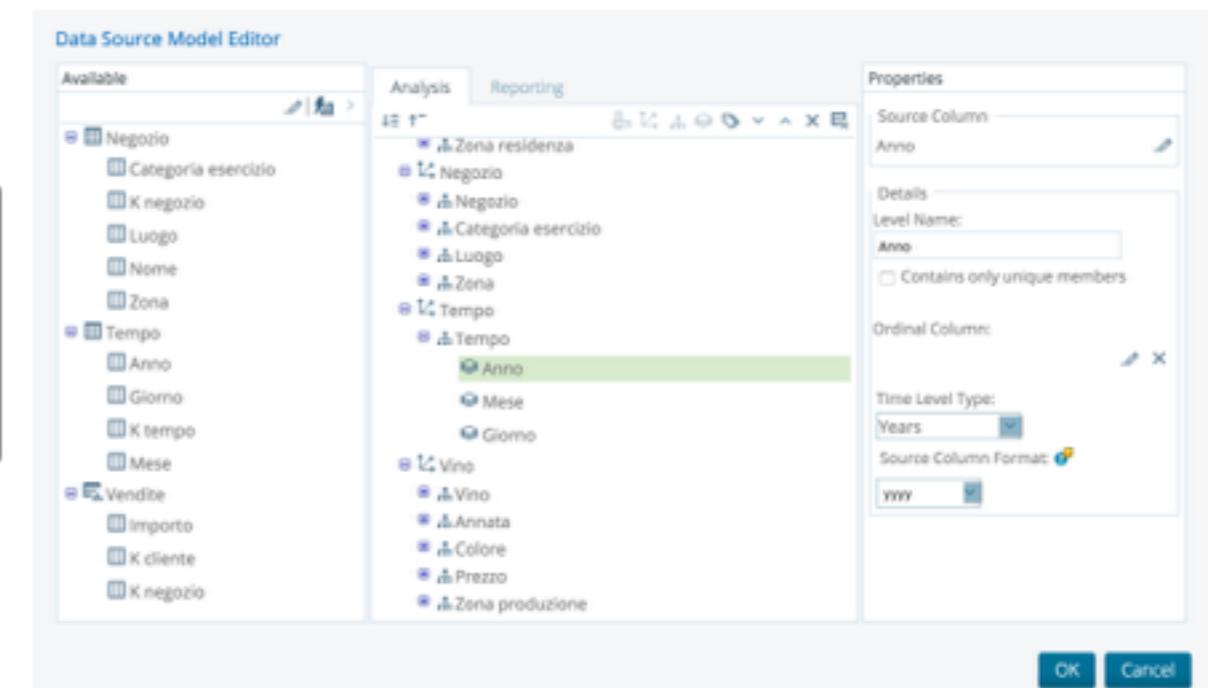
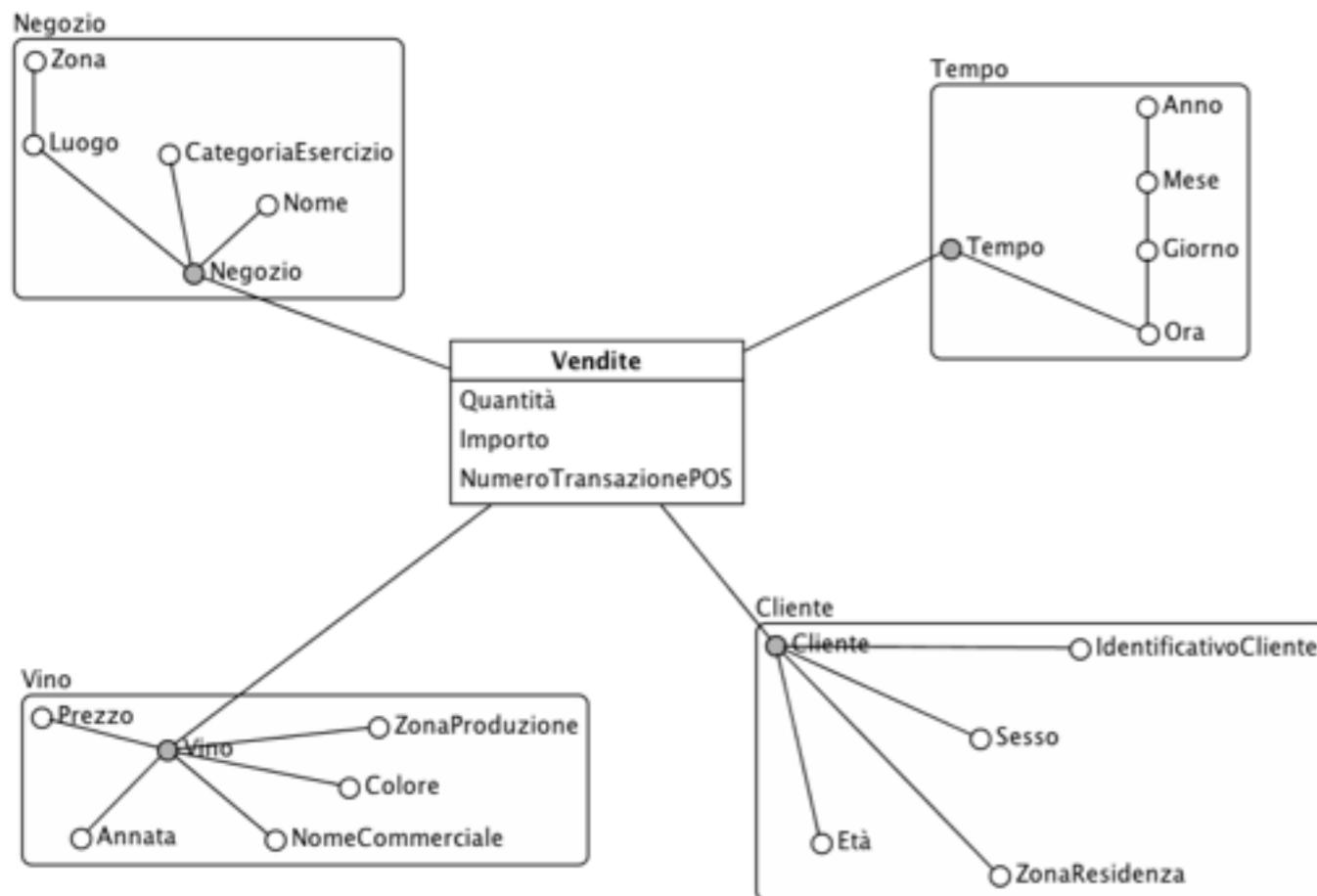


- [Dimensione.NomeGerarchia]
- Esempi
  - [Tempo] : gerarchia di default
  - [Negozio.Categoria esercizio]
  - [Vino.Colore]
  - [Vino] : gerarchia di default



# Meta-definizioni

## Nome di livello di gerarchia

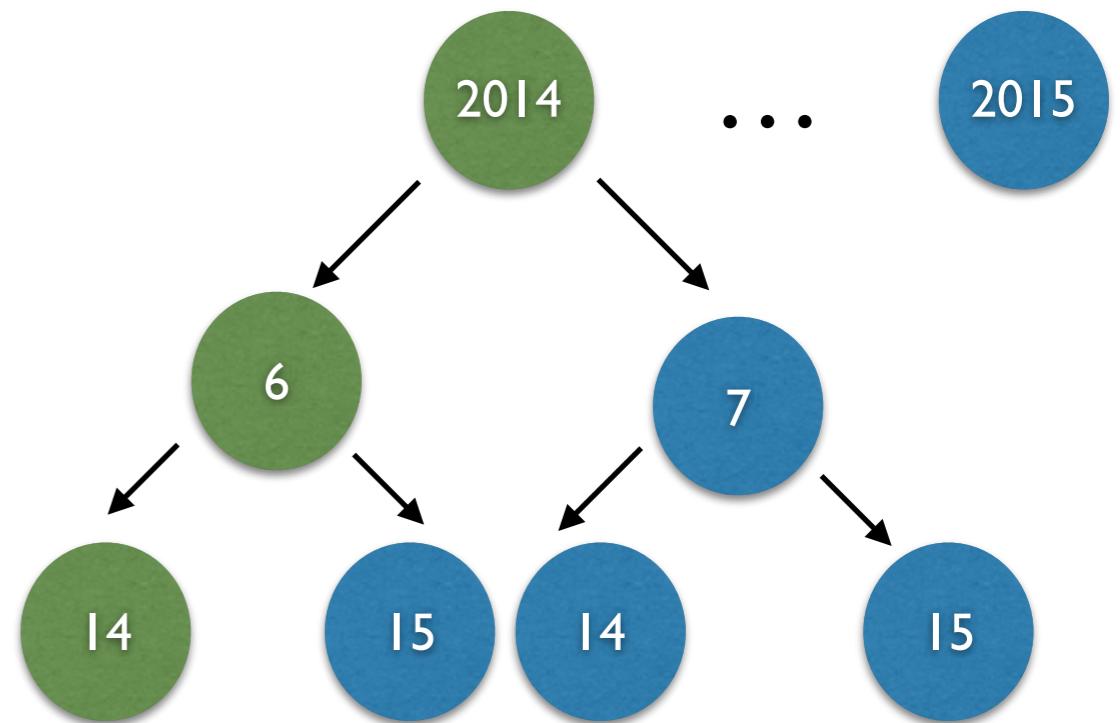


- Scelta di un livello di una gerarchia.
  - [Dimensione.Gerarchia].[Livello]
- Esempi
  - [Data].[Anno]
  - [Data].[Mese]
  - [Data].[Giorno]

# Definizioni

## Nome di un membro di un livello (I)

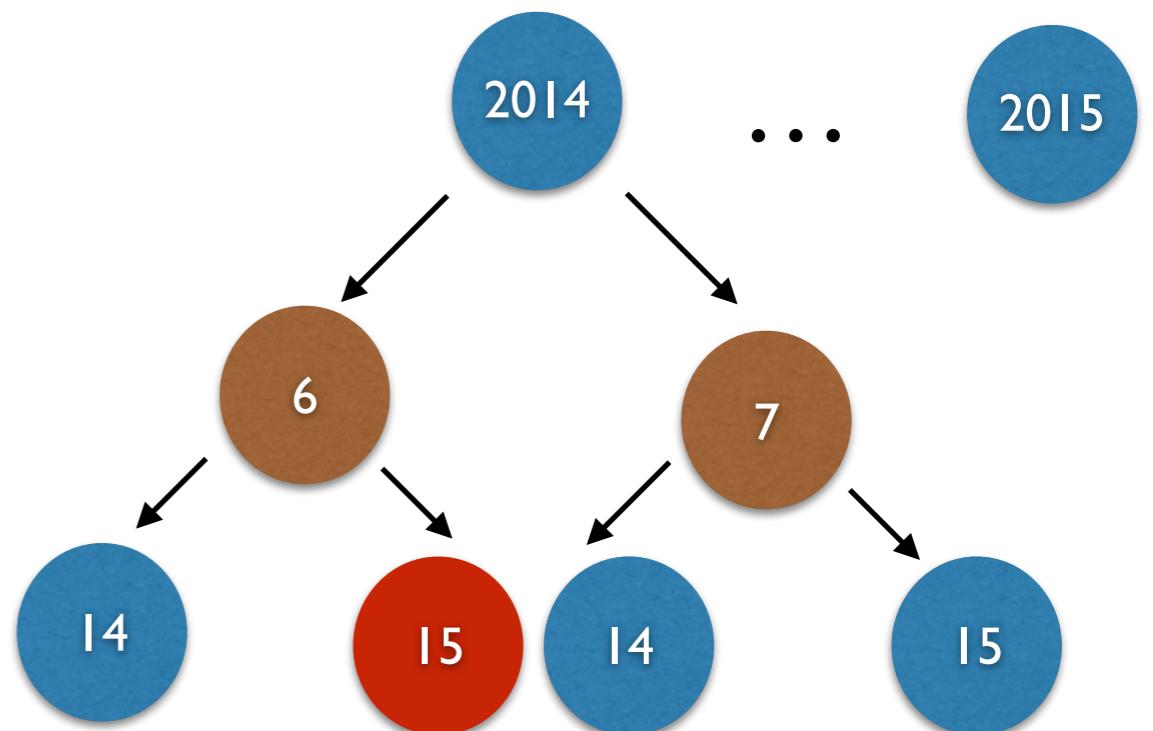
- Il **valore** di un elemento (**membro**) vero e proprio di una dimensione.
- A qualunque livello nella gerarchia, è il **valore dell'attributo identificativo del livello**.
- Membri di livelli più granulari possono essere specificati a partire da membri di livelli meno granulari.
- Si specificano cammini nell'albero della gerarchia dai livelli meno granulari a quelli più granulari.
  - [Dimensione.Gerarchia].[Membro I]
  - [Dimensione.Gerarchia].[Membro I].[Membro 2]
  - [Dimensione.Gerarchia].[Livello I].[Membro I]
  - [Dim.Gerarchia].[Livello I].[Membro I].[Membro 2]
- Esempi:
  - [Data].[2014]
  - [Data].[2014].[6]
  - [Data].[2014].[6].[14]



# Definizioni

## Nome di un membro di un livello (2)

- Non si possono specificare cammini che partono da un livello maggiore del primo e comprendono un nome di membro.
- L'idea (probabilmente) è non avere cammini a granularità variabile.
  - [Data].[Anno].[2014]
  - [Data].[2014].[6]
  - [Data].[Anno].[2014].[6].[14]
  - Sono esempi validi.
  - **[Data].[Mese].[15]** non è valido.
- In un cammino, il nome di un membro di un livello può essere seguito solo da il nome di un membro di un livello più granulare o da una **funzione speciale sui membri**.
- In un cammino, il nome di un livello può essere seguito da il nome di uno o più membri o da una **funzione speciale sui livelli**.



# Misure

- Esiste una dimensione speciale denominata Measures.
- Si può referenziare quindi con [Measures].
- Ha un'unica gerarchia di default con un solo livello.
- I membri dell'unico livello assumono il nome delle misure del cubo di input.
- [Measures].[Importo]
- [Measures].[Quantita]

# [something]

- **[something.this].[what is this].[what is that]**
- **something**
  - nome di una dimensione.
- **this**
  - nome di una gerarchia
  - può essere omesso per indicare quella di default.
- **what is this**
  - nome di un livello della gerarchia this
  - nome di un membro del primo livello della gerarchia this
- **what is that**
  - nome di un membro del secondo livello della gerarchia this

# Tupla

- Una lista di espressioni che denotano membri.
  - (**[Negozio.Categoria esercizio].[alimentari]** , **[Cliente.Zona residenza].[BE]**)
- I membri di una tupla possono appartenere alla stessa dimensione, ma non alla stessa gerarchia.
- (**[Negozio.Categoria esercizio].[alimentari]**,**[Negozio.Luogo].[4042 Ipsum Rd.]**)
  - Valida: appartengono a due gerarchie diverse della dimensione Negozio.
- (**[Negozio.Categoria esercizio].[alimentari]**,**[Negozio.Categoria esercizio].[bar]**)
  - Non valida: entrambi i membri sono della gerarchia Categoria esercizio.
- Una tupla rappresenta un punto su un asse del cubo di output.
  - Combinazione di membri di dimensioni dell'input.

# Set ( | )

- Un insieme di tuple.
- Può essere specificato per enumerazione delle tuple o attraverso funzioni speciali.
- {[Negozio.Categoria esercizio].[alimentari]}
  - L'insieme che contiene la sola tupla relativa al membro alimentari.
- {[Negozio.Categoria esercizio].[alimentari]}, {[Negozio.Categoria esercizio].[bar]}
  - L'insieme che contiene due tuple relative a negozi di alimentari e bar.
- {[Negozio.Categoria esercizio].[alimentari],[Tempo].[2000]}, {[Negozio.Categoria esercizio].[alimentari],[Tempo].[2001]}
  - L'insieme che contiene due tuple relative ai negozi di alimentari, negli anni 2000 e 2010
- Un set rappresenta tutti i punti di un asse del cubo di output.
- Nell'input rappresenta un insieme di celle o un insieme di iper-piani.

# Set (2)

- Sono ammessi solo Set le cui tuple hanno membri ordinatamente delle stesse gerarchie, anche a livelli di granularità diversi.
- Ad esempio:
- $\{([Negozi.Categoria\ esercizio].[alimentari],[Tempo].[2000]),([Negozi.Categoria\ esercizio].[alimentari],[Tempo].[2001])\}$
- ma non:
  - $\{([Negozi.Categoria\ esercizio].[alimentari],[Tempo].[2000]),([Negozi.Luogo],[Tempo].[2001])\}$
- Anche valida (livelli di granularità diversi):
  - $\{([Negozi.Categoria\ esercizio].[alimentari],[Tempo].[2000]),([Negozi.Categoria\ esercizio].[alimentari],[Tempo].[2001].[5])\}$
- I Set sono in realtà sottoinsiemi del prodotto cartesiano di gerarchie diverse.
- Sono in realtà multiset perché ammettono i duplicati.

# Una query

```
SELECT
    <<set>> | <<tuple>> | <<member>> on AXIS 1,
    <<set>> | <<tuple>> | <<member>> on AXIS 2, ...,
    <<set>> | <<tuple>> | <<member>> on AXIS n,

    FROM [<<fact table>>]
    WHERE <<set>> | <<tuple>> || <<member>>
```

- La clausola **SELECT** specifica quali Set sono mappati su quali assi.
  - Ogni tupla di ogni set rappresenta un punto sul relativo asse.
  - Si può usare una tupla se il set ne ha una sola.
    - Si può usare un membro se la tupla ne ha uno solo.
- **FROM** identifica il cubo sorgente.
- **WHERE** effettua lo slicing selezionando quale parte del cubo di input mappare nella clausola **SELECT**.
  - Si possono identificare celle (con tuple complete)
  - Iper-piani (con tuple incomplete o membri singoli)
  - Insiemi di celle o di iper-piani (set)
- **SELECT** e **WHERE** devono riguardare gerarchie diverse.
- L'output è ottenuto **aggregando** i valori delle misure per tutte le intersezioni degli assi.

# Query con due assi

```
SELECT
    <<set>> on ROWS,
    <<set>> on COLUMNS

FROM [<<fact table>>]
WHERE <<set>> | <<tuple>> || <<member>>
```

- Tipicamente si utilizzano solo due assi nel cubo di output.
  - Rappresentazione bidimensionale
  - Gli assi si chiamano ROWS e COLUMNS.

# Query MDX

Importo e quantità venduta per il vino “cometa”

```
select NON EMPTY { [Measures].[Importo], [Measures].[Quantita] }
ON COLUMNS,
NON EMPTY {([Vino].[cometa])}
ON ROWS
from [DHW_FINAL]
```

	Measures	
Vino	• Importo	• Quantita
cometa	166090	9770

# Query MDX

Importo e quantità venduta per il vino “cometa” e per il vino “frappato”

```
select NON EMPTY { [Measures].[Importo], [Measures].[Quantita] }
ON COLUMNS,
NON EMPTY {([Vino].[cometa]), ([Vino].[frappato])}
ON ROWS
from [DHW_FINAL]
```

	Measures	
Vino	• Importo	• Quantita
cometa	166090	9770
frappato	67277	9611

# Funzioni speciali

- **MEMBERS**
  - permette di generare l'insieme di tutti i membri di un livello.
- **NON EMPTY**
  - permette di scartare nell'output i punti degli assi per i quali non ci sono valori.

# Query MDX

Quantità venduta per annata del vino “cometa”

```
select NON EMPTY {[Measures].[Quantita]} ON COLUMNS,  
NON EMPTY {[Vino.Annata].members} ON ROWS  
from [DHW_FINAL]  
where {[Vino].[cometa]}
```

	Measures
Annata	● Quantita
All Vino.Annatas	9770
2010	9770

# Query MDX

Importo totale per colore del vino (righe) e per categoria del negozio (colonne).

```
select NON EMPTY {[Negozio.Categoria esercizio].members} ON  
COLUMNS,  
NON EMPTY {[Vino.Colore].members} ON ROWS  
from [DHW_FINAL]
```

	Categoria esercizio	alimentari	bar	supermercato
Colore	All Negozio.Categoria esercizios			
All Vino.Colores	951533	285001	284625	381907
bianco	277862	82239	83781	111842
giallo	377357	114673	111617	151067
rosa	70077	21259	20363	28455
rosso	226237	66830	68864	90543

# Query MDX

Importo per fasce di età e nome del vino.

```
select  
    NON EMPTY {[Cliente.Eta].members} ON COLUMNS,  
    NON EMPTY {[Vino].members} ON ROWS  
from [DHW_FINAL]
```

		Eta																	
Vino	All Cliente.Etas	19	30	34	36	37	42	44	49	51	58	60	63	70	74	76			
All Vinos	951533	47717	47853	93654	48546	47357	47107	48421	45641	94799	191124	47398	47227	48479	49781	46429			
alastro	71253	3640	3395	7210	3395	3465	3486	3388	3423	7448	13797	3703	3402	3913	3934	3654			
cometa	166090	8670	7990	15470	8942	7497	8024	8738	8228	17459	33218	8109	8449	7939	8908	8449			
eruzione 1614 carricante	140014	6678	7392	13930	7378	7392	7588	7504	6538	13650	27636	6580	7098	7784	7196	5670			
eruzione 1614 nerello mascalese	158960	8016	7968	16640	7712	8656	7840	8000	7952	15056	33536	7408	6960	7968	8224	7024			
etna bianco	228022	11617	11294	22350	11401	11440	11010	11572	10125	22353	45914	12299	11983	11126	11712	11826			
frappato	67277	3171	3514	6419	3717	3325	3220	3241	3416	6916	13412	3381	3122	3409	3577	3437			
la segreta	49840	2355	2730	5090	2340	2180	2530	2415	2410	4700	10220	2355	2510	2805	2450	2750			
rosé	70077	3570	3570	6545	3661	3402	3409	3563	3549	7217	13391	3563	3703	3535	3780	3619			

# Query MDX

Importi per fascia d'età e zona di residenza del cliente.

```
select  
NON EMPTY {[Cliente.Zona residenza].members} ON COLUMNS,  
NON EMPTY {[Cliente.Eta].members}) ON ROWS  
from [DHW_FINAL]
```

	Zona residenza	BE	Berlin	CAM	Ceará	Comunitat Valenciana	KA	KN	MS	Madrid	Manitoba	Mississippi	NI	Noord Brabant	ON	OV	PA	RM	SP	South Island	U	
Eta	All Cliente.Zona residencias	951533	49781	47398	46357	47107	47357	47084	48442	48546	45641	46429	46568	47952	47086	48479	47717	48421	48161	47853	47227	47927
19		47717														47717						
30		47853																		47853		
34		93654											46568		47086							
36		48546								48546												
37		47357					47357															
42		47107		47107																		
44		48421															48421					
49		45641							45641													
51		94799		46357				48442														
58		191124					47084					47952					48161			47927		
60		47398	47398																			
63		47227																			47227	
70		48479													48479							
74		49781	49781																			
76		46429							46429													

# Funzioni speciali

- **crossjoin** (set 1, set 2)
  - prende in input due insiemi e genera l'insieme prodotto cartesiano delle tuple.
  - Utile per combinare su un asse gli elementi di più dimensioni.

# Query MDX

Importi per colore del vino, zona di residenza del cliente, zona del negozio, con vendite nel 2000.

```
select  
NON EMPTY {[Vino.Colore].members} ON COLUMNS,  
NON EMPTY crossjoin({([Cliente.Zona residenza].members)},  
{([Negozio.Luogo].members)}) ON ROWS  
from [DHW_FINAL]  
where [Tempo].[2000]
```

# Query MDX

		Colore					
Zona residenza	Luogo	All Vino.Colores	bianco	giallo	rosa	rosso	
All Cliente.Zona residenzas	All Negozio.Luogos	96040	27969	38273	7000	22798	
	3560 Congue Street	9164	2392	4064	581	2127	
	4042 Ipsum Rd.	10139	2809	4139	805	2386	
	4571 Accumsan St.	9417	2859	3625	749	2184	
	4586 Pede Rd.	9348	2771	3481	861	2235	
	9451 Iaculis Street	9654	3025	3509	805	2315	
	Ap #701-5553 Iaculis, Rd.	9649	2947	3712	672	2318	
	Ap #838-2014 Quis, Rd.	9519	2853	3876	707	2083	
	P.O. Box 450, 1030 Volutpat. St.	9642	2935	3797	756	2154	
	P.O. Box 486, 6366 Molestie St.	10242	2938	4124	518	2662	
BE	P.O. Box 601, 3107 Elementum Road	9266	2440	3946	546	2334	
	All Negozio.Luogos	5236	1684	2105	336	1111	
	3560 Congue Street	549	112	179	35	223	
	4042 Ipsum Rd.	527	201	247	7	72	
	4571 Accumsan St.	616	204	193	70	149	
	4586 Pede Rd.	394	197	139	42	16	
	9451 Iaculis Street	372	154	104	35	79	
	Ap #701-5553 Iaculis, Rd.	401	144	215	14	28	
	Ap #838-2014 Quis, Rd.	572	170	244	63	95	

# Query MDX

Quantità vendute e importi totali in tutti i mesi dell'anno per vino.

```
select  
{([Measures].[Quantita]), ([Measures].[Importo])} ON ROWS,  
NON EMPTY crossjoin({([Tempo].[mese].members)},  
{([Vino].members)}) ON COLUMNS  
from [DHW_FINAL]
```

# Query MDX

Tempo															Vino														
3															5														
Vino															Vino														
Measures	All Vinos	alastro	cometa	eruzione 1614	eruzione 1614	etna bianco	frappato	la segreta	rosé	santa cecilia	All Vinos	alastro	cometa	eruzione 1614	eruzione 1614	etna bianco	frappato	carricante	nero	d'avola									
Quantita	10001	1042	965	1041	983	1987	1010	1018	1000	955	9957	999	998	1036	1057	1848	983												
Importo	96040	7294	16405	14574	15728	22879	7070	5090	7000		95815	6993	16966	14504	16912	21258	6881												

# Query MDX

Quantità di vino (indipendentemente dal vino specifico) venduto; risultati per zona del negozio. Interessa solo l'anno 2000.

```
select  
{([Measures].[Fact Count])} ON COLUMNS,  
NON EMPTY [Negozio.Luogo].MEMBERS on ROWS  
from [DHW_FINAL]  
where [Tempo].[Anno].[2000]
```

# Query MDX

	Measures
Luogo	<input checked="" type="radio"/> Fact Count
All Negozio.Luogos	2.000
3560 Congue Street	200
4042 Ipsum Rd.	200
4571 Accumsan St.	200
4586 Pede Rd.	200
9451 Iaculis Street	200
Ap #701-5553 Iaculis, Rd.	200
Ap #838-2014 Quis, Rd.	200
P.O. Box 450, 1030 Volutpat. St.	200
P.O. Box 486, 6366 Molestie St.	200
P.O. Box 601, 3107 Elementum Road	200

# Misure derivate

```
WITH MEMBER [Measures].[Importo Medio] as  
[Measures].[Importo] / [Measures].[Quantita]
```

- Si possono combinare le misure esistenti tra loro o con numerici per calcolare delle misure derivate.
- Si usa una tecnica analoga per estrarre i valori degli attributi dei membri.

```
WITH MEMBER [Measures].[Nome misura] as  
[Dimensione.Gerarchia].CurrentMember.Properties("NomeAttributo")
```

# Query MDX

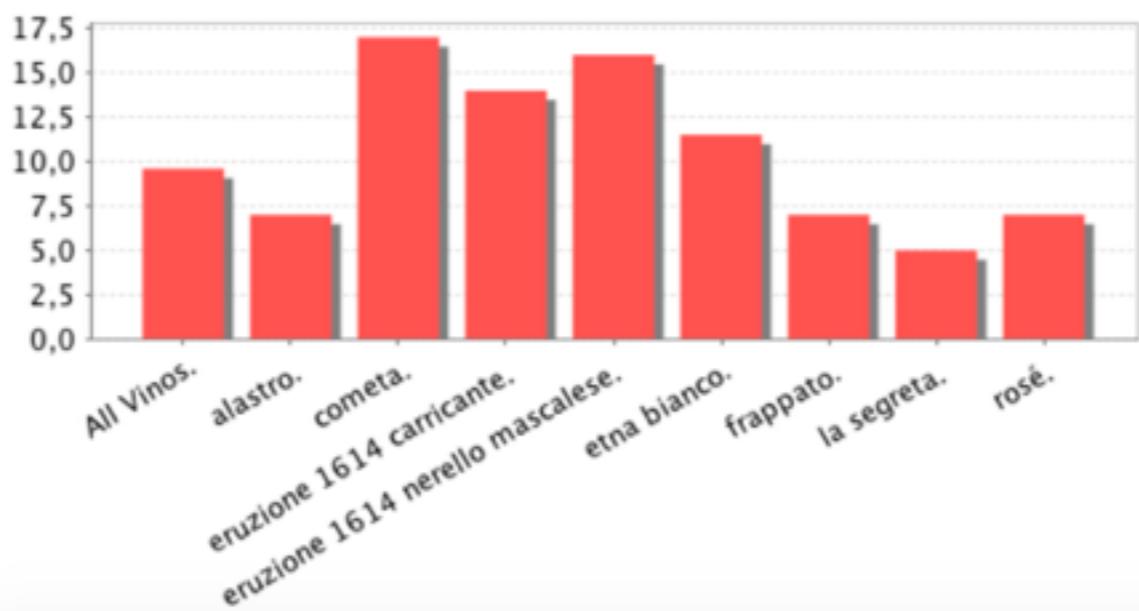
Importo medio pagato per vino.

```
with member [Measures].[Importo Medio] as [Measures].[Importo] /  
[Measures].[Quantita]  
  
select  
{([Measures].[Importo Medio])} ON COLUMNS,  
NON EMPTY [Vino].MEMBERS on ROWS  
  
from [DHW_FINAL]
```

# Query MDX

Vino	Measures
Vino	• Importo Medio
All Vinos	10
alastro	7
cometa	17
eruzione 1614 carricante	14
eruzione 1614 nerello mascalese	16
etna bianco	12
frappato	7
la segreta	5
rosé	7

Slicer:



# Funzioni utili

- **children**
  - Ritorna tutti i membri figli di un membro dato.
  - **Members** ritorna tutti gli elementi di un dato livello e un elemento artificiale “All <>s” che contiene i totali.
- molti altri operatori. Ne vedremo alcuni.
- <http://mondrian.pentaho.com/documentation/mdx.php>

# Query MDX

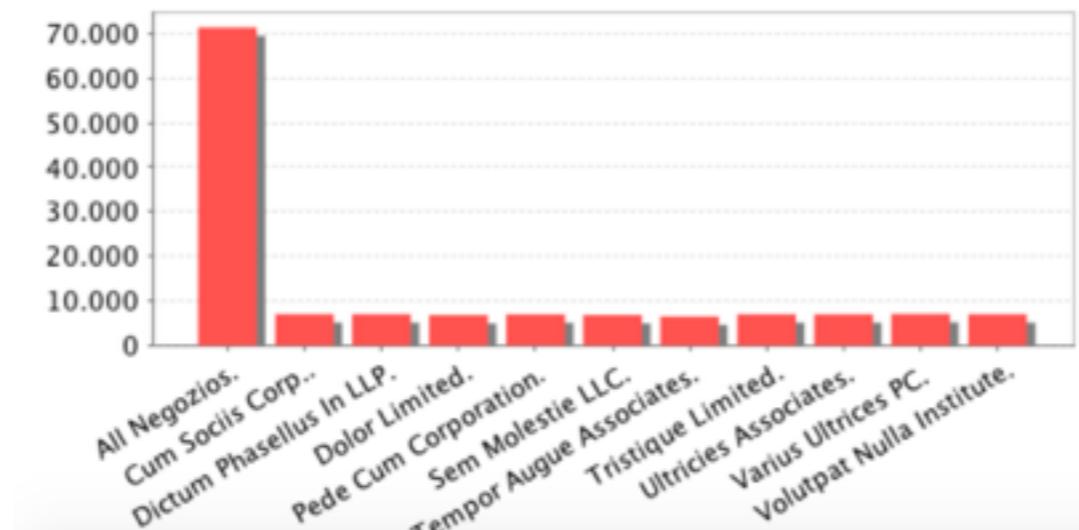
Gli importi delle vendite relative ai vini prodotti a Ulmo,  
suddivise per negozio.

```
select  
{([Measures].[Importo])} ON COLUMNS,  
NON EMPTY [Negozio].members on ROWS  
from [DHW_FINAL]  
where [Vino.Zona produzione].[ulmo]
```

# Query MDX

Measures	
Negozi	• Importo
All Negozios	71253
Cum Sociis Corp.	7224
Dictum Phasellus In LLP	7210
Dolor Limited	7021
Pede Cum Corporation	7175
Sem Molestie LLC	7028
Tempor Augue Associates	6713
Tristique Limited	7203
Ultricies Associates	7196
Varius Ultrices PC	7273
Volutpat Nulla Institute	7210

Slicer: [Zona produzione=ulmo]



# Dashboards



# Costruire dashboard

- Realizziamo dei cruscotti che diano una visione sintetica per un certo utente
- Un cruscotto è un insieme di componenti grafici che mostrano risultati di interrogazioni
- Le interrogazioni sono effettuate sul nostro data mart
- Dashboard da includere in siti web o applicazioni enterprise

# Strumento CDE

Screenshot of the Pentaho Business Analytics CDE (Common Data Environment) interface.

The top navigation bar includes: File, View, Tools, Help, Home, and admin.

The main menu on the left contains: Browse Files, Create New, Manage Data Sources, Documentation, Jpivot View, CDE Dashboard (which is highlighted with a red arrow), and Data Source.

The center features the "Pentaho Business Analytics" logo and a call-to-action message: "and contribute with knowledge. Some ways how to do it."

A diagram illustrates a network of data flow or connectivity between various components.

On the right, there are links to Documentation (with a document icon), Forums (with a speech bubble icon), Mailing Lists (with an envelope icon), and Blogs (with a pen icon).

The "Recents" section lists: test\_dashboard.c... (with a star icon) and test\_dashboard.cda (with a star icon).

The "Favorites" section shows: You haven't.

# Strumento CDE

The screenshot shows the CDE Dashboard interface. At the top, there's a toolbar with 'CDE' logo, 'New', 'Save', 'Save as...', 'Reload', and 'Settings' buttons, along with a 'New Dashboard' button. Below the toolbar is a 'Layout Structure' panel with icons for creating various dashboard components like cards, charts, and tables. To its right is a 'Properties' panel with columns for 'Property' and 'Value'. On the far right of the interface, there's a toolbar with icons for previewing, saving, and other operations. Red arrows point from the text labels 'Layout', 'Data Source', and 'Componenti' to the corresponding sections of the interface.

- 1. ideare il layout
- 2. realizzare il layout
- 3. creare i data source
- 4. creare i componenti
- 5. legare i componenti al layout

# Ideare il layout

- Vogliamo realizzare un cruscotto con quattro grafici
  - istogramma delle vendite totali (importi) suddivise per fascia di età dei clienti
  - line chart delle quantità totali vendute negli anni
  - pie chart delle vendite totali (importi) per tipo di vino
  - tabella dei dati delle vendite per area geografica dei negozi
- La dashboard deve avere un titolo e quattro riquadri con i vari grafici

# Layout (I)

CDE Dashboard (2) ×

**CDE** New Save Save as... Reload Settings New Dashboard \*

Layout Structure

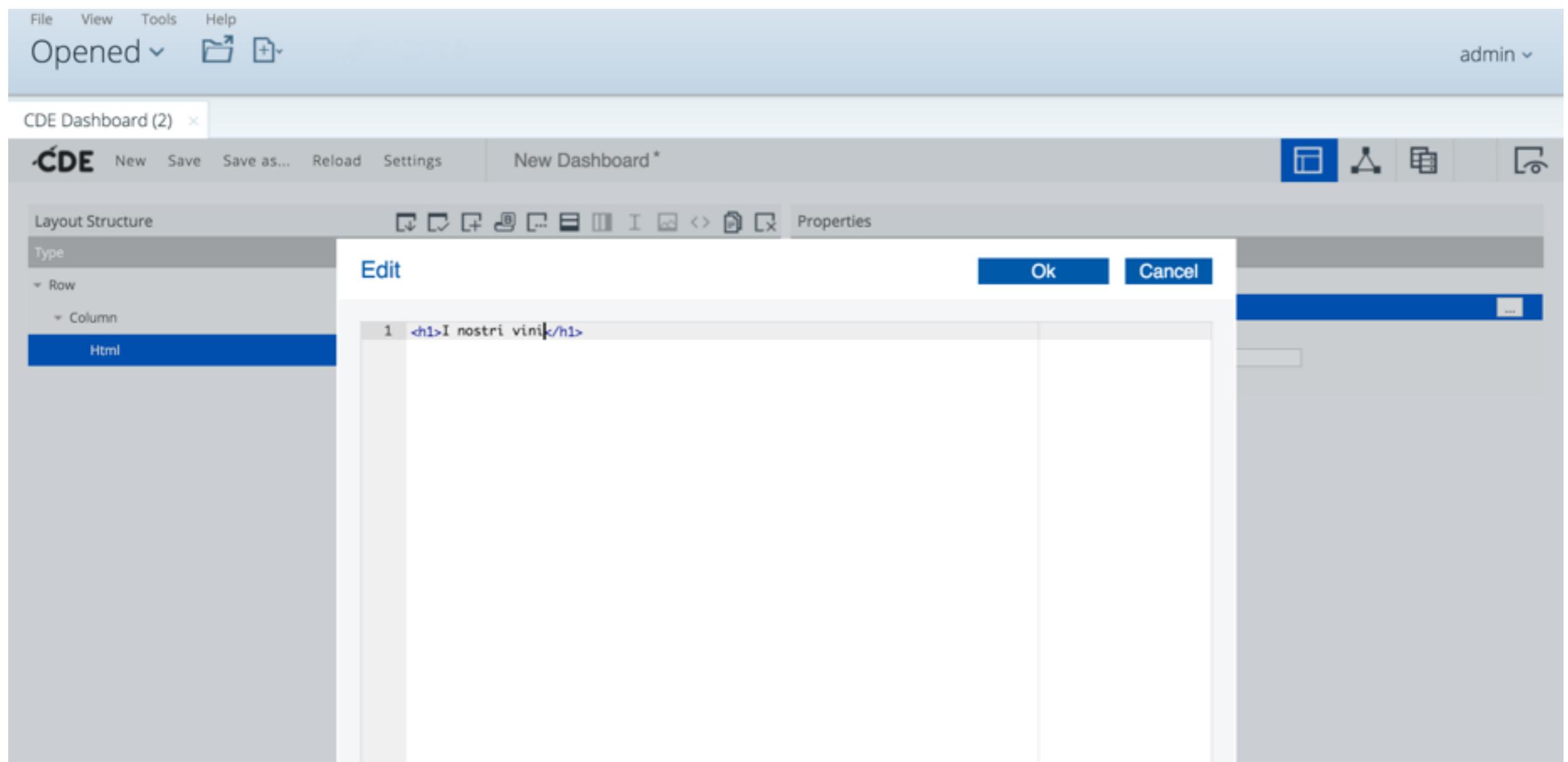
Type	Name
Row	Intestazione
Column	Titolo

Properties

Property	Value
Name	Titolo
Extra Small Devices	12
Small Devices	-
Medium Devices	-
Large Devices	-
Bootstrap Css Class	-
Height	-
BackgroundColor	<input type="color"/>
Corners	Simple
Text Align	Center
Css Class	-

Icons:

# Layout (2)



# Layout (3)

Layout Structure

The screenshot shows a layout editor interface with two main sections: 'Layout Structure' on the left and 'Properties' on the right.

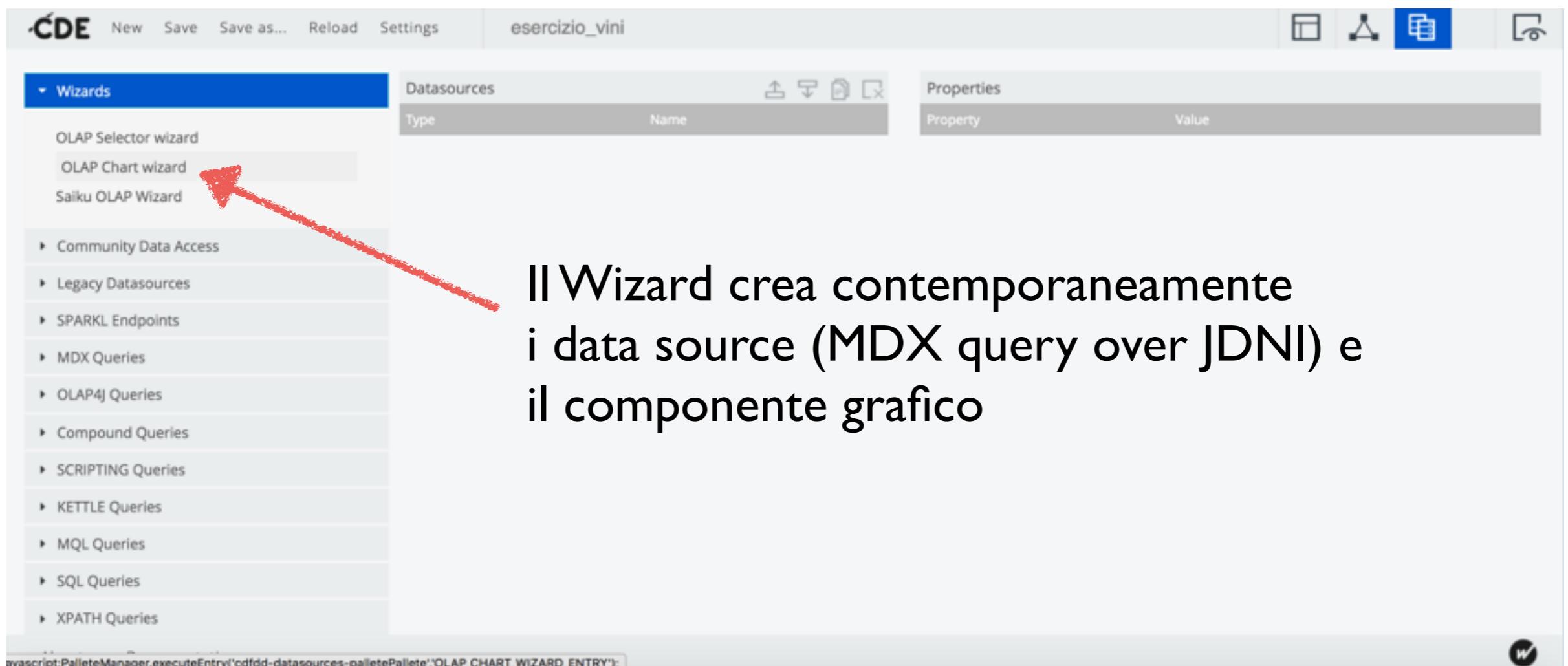
**Layout Structure:**

- Type: Row
  - Name: Intestazione
- Type: Column
  - Name: Titolo
- Type: Row
  - Column: Colonna1
  - Column: Colonna2
- Type: Row
  - Column: Colonna3
  - Column: Colonna4

**Properties:**

Property	Value
Name	Colonna4
Extra Small Devices	12
Small Devices	-
Medium Devices	-
Large Devices	-
Bootstrap Css Class	-
Height	-
BackgroundColor	<input type="color"/>
Corners	Round
Text Align	Center
Css Class	-

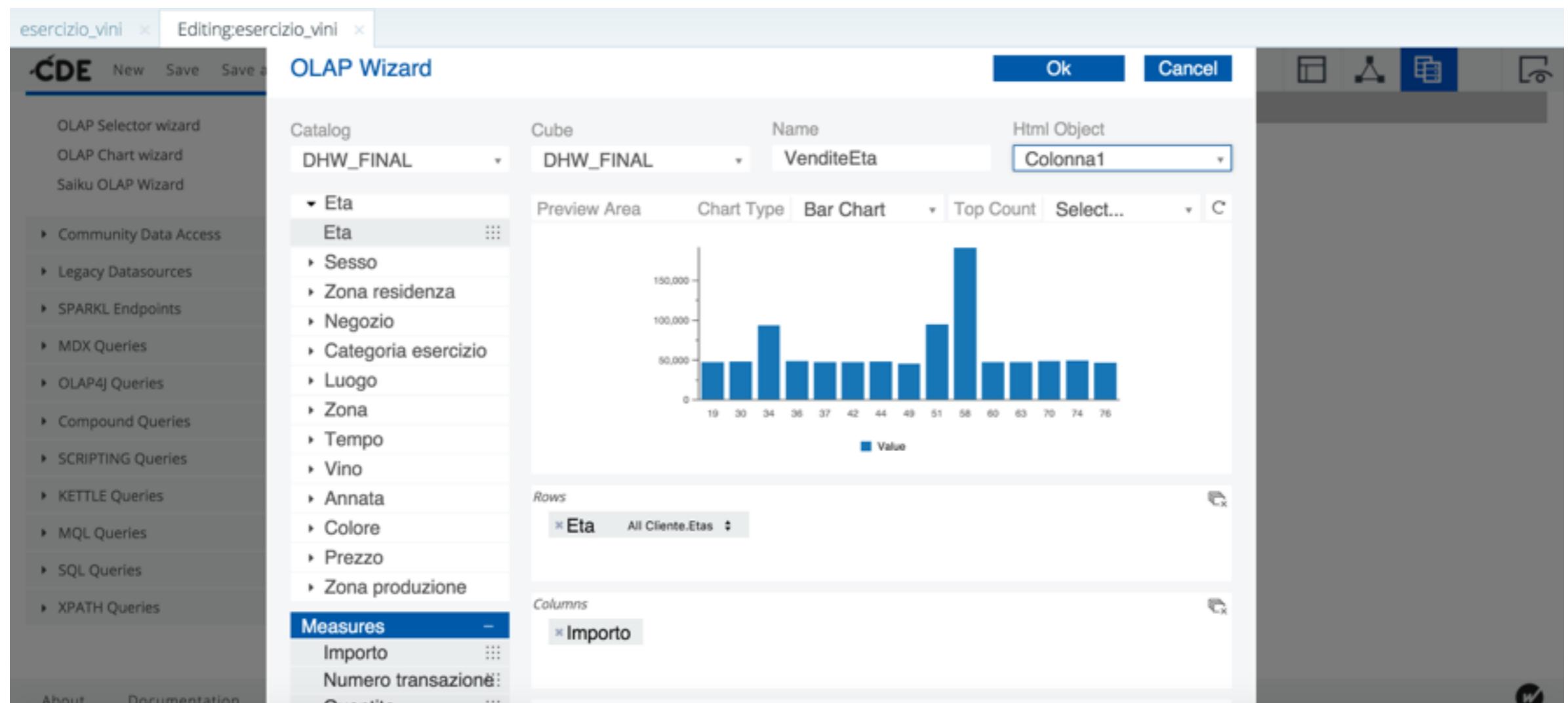
# Data source



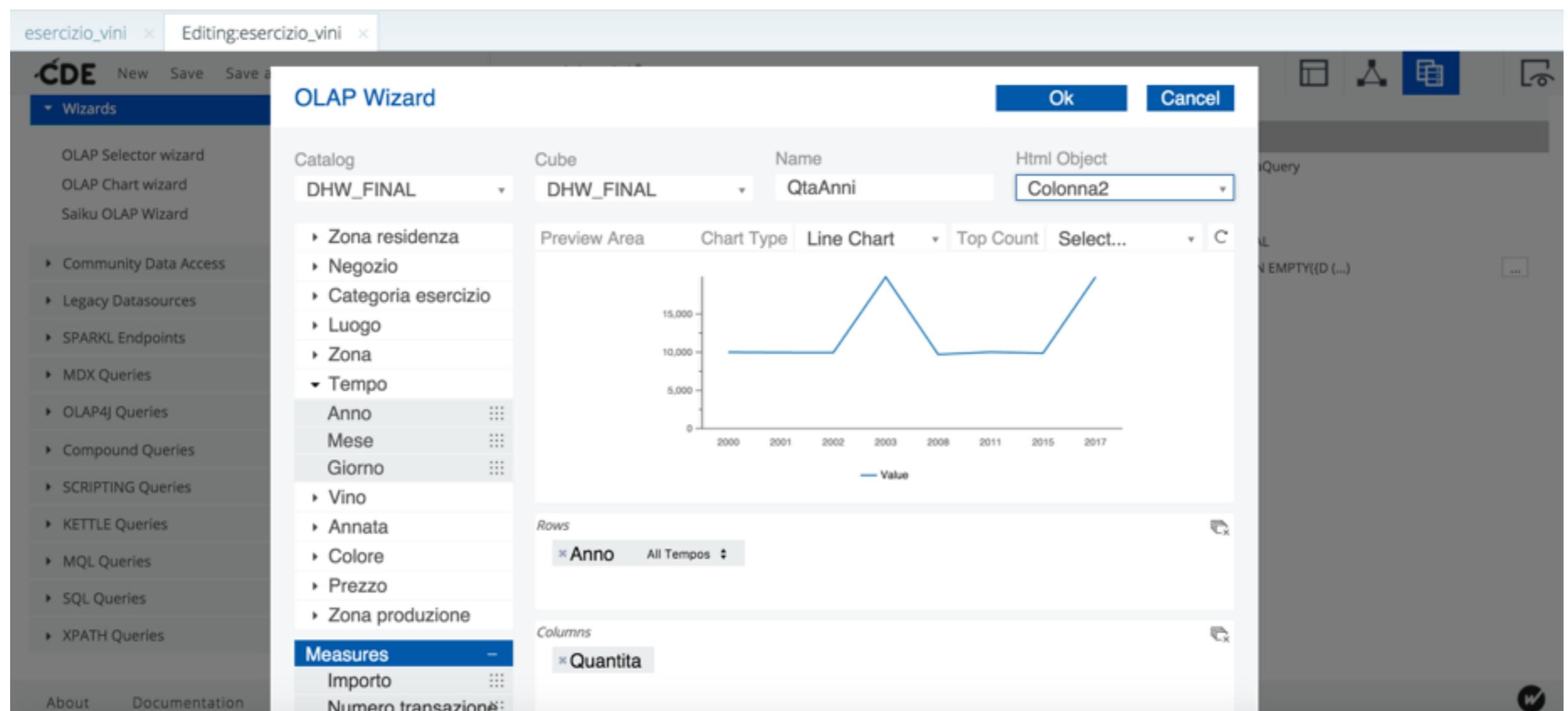
The screenshot shows the CDE (Community Data Engine) interface. The top navigation bar includes 'New', 'Save', 'Save as...', 'Reload', and 'Settings'. The title bar says 'esercizio\_vini'. The main area has three tabs: 'Wizards' (selected), 'Datasources', and 'Properties'. The 'Wizards' tab contains a list of wizards, with 'OLAP Chart wizard' highlighted and a red arrow pointing to it from the text below. The 'Datasources' and 'Properties' tabs are empty. The bottom status bar shows the JavaScript code: 'avascript:PaletteManager.executeEntry('cdfdd-datasources-palletePallete','OLAP\_CHART\_WIZARD\_ENTRY');' and a small WPS Office logo.

Il Wizard crea contemporaneamente i data source (MDX query over JDNI) e il componente grafico

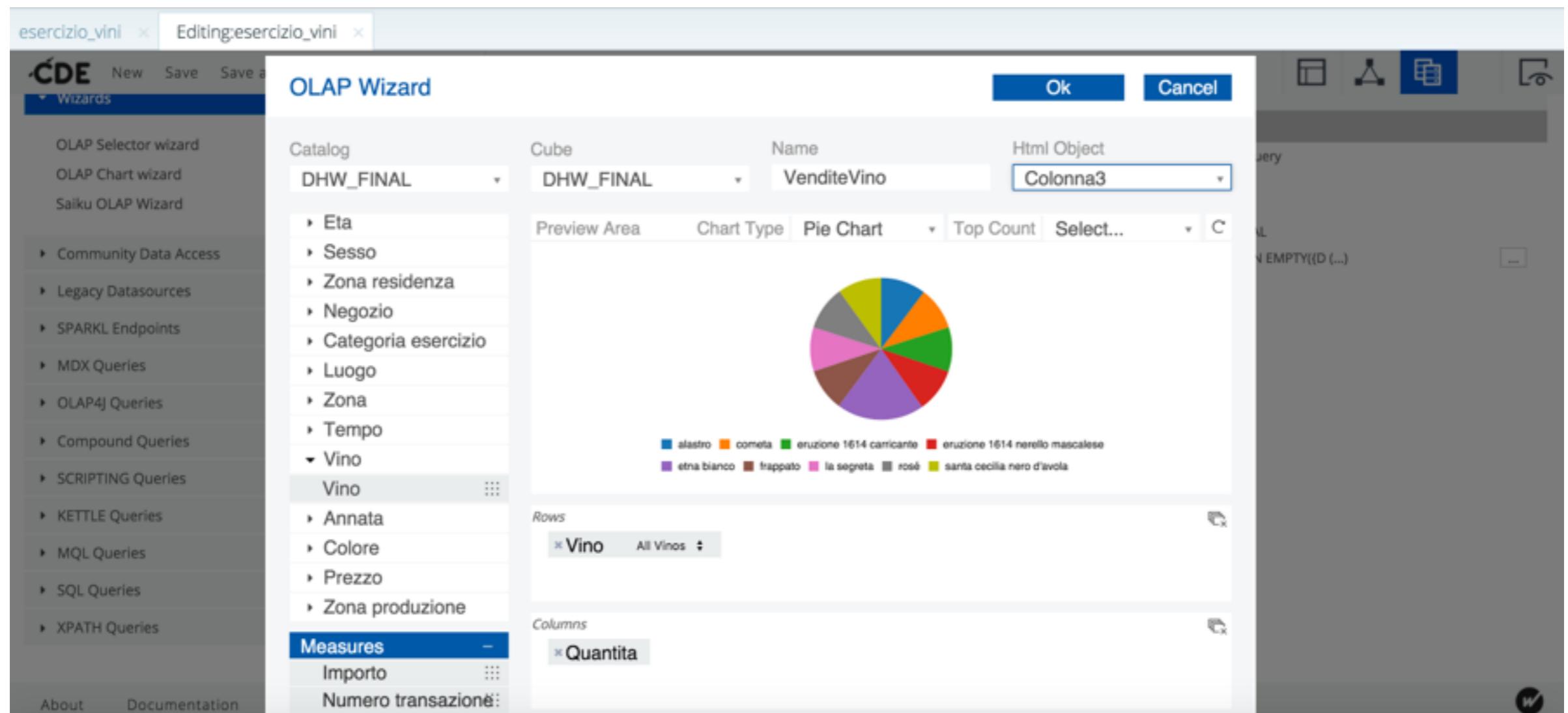
# Vendite per fascia d'età



# Quantità vendute negli anni



# Vendite per tipo di vino



# Vendite per area geografica del negozio (d.s.)

The screenshot shows the CDE (Community Data Engine) interface. The title bar says "esercizio\_vini" and "Editing:esercizio\_vini". The menu bar includes "New", "Save", "Save as...", "Reload", and "Settings". The main window has a sidebar with "Wizards", "Community Data Access", "Legacy Datasources", "SPARKL Endpoints", and "MDX Queries" (which is selected). The "MDX Editor" tab is active, showing the following query:

```
Parameters +  
1 select NON EMPTY {[Negozio.Luogo].members} ON COLUMNS,  
2     NON EMPTY {[Measures].[Importo]} ON ROWS  
3     from [DHW_FINAL]
```

The "Ok" and "Cancel" buttons are visible at the top right of the editor.

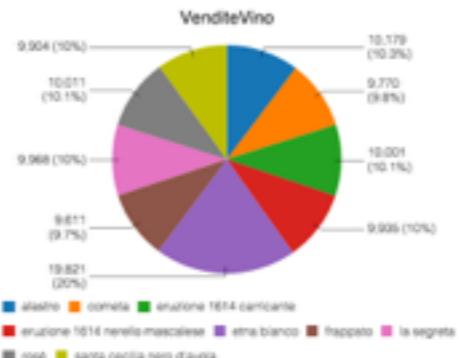
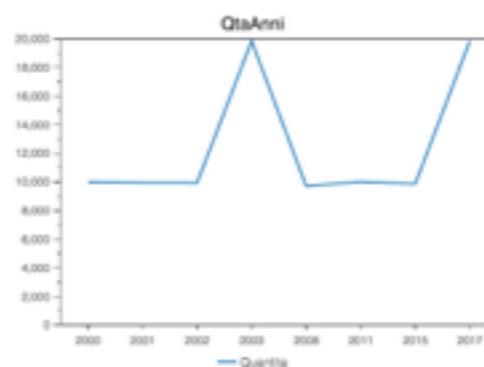
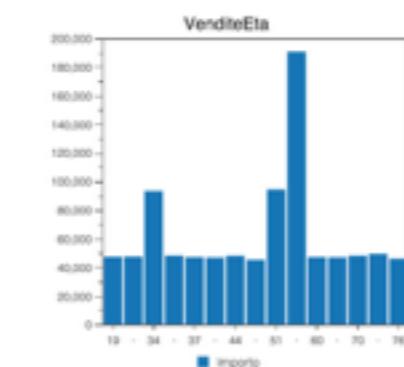
# Vendita per area geografica del negozio (component)

The screenshot shows a user interface for managing components, likely in a reporting or analytical tool. On the left, a sidebar lists categories: 'Charts' and 'Others'. Under 'Others', several components are listed: Analyzer Component, Button Component, Comments component, Execute Analyzer Component, Execute Prpt Component, Execute XAction Component, Freeform Component, Pivot Component, Navigation Menu Component, Pivot Link Component, PRPT Component, Query Component, Schedule Prpt Component, Table Component, Template Component, and Text Component. A blue bar highlights the 'Others' category. In the center, a table titled 'Components' lists items under 'Type' (e.g., Group, CCC Bar Chart, CCC Line Chart, CCC Pie Chart) and 'Name' (e.g., Charts, VenditeEtaChart, QtaAnniChart, VenditeVinoChart, Others). A blue bar highlights the 'table Component' row, which has the name 'VenditeAreaNegozi'. To the right, a 'Properties / Advanced Properties' panel shows settings for the selected component:

Property	Value
Name	VenditeAreaNegozi
Listeners	
Column Headers	
Column Types	
Parameters	
Datasource	VenditeAreaNegozi
HtmlObject	Colonna4
clickAction	

# La nostra dashboard

## I nostri vini



Show 10 entries Search:

Negozi.Luogo	Importo
All Negozi.Lugos	951533
3560 Congue Street	92981
4042 Ipsum Rd.	95261
4571 Accumsan St.	93613
4586 Pede Rd.	95609
9451 Iaculis Street	96269
Ap #701-5553 Iaculis, Rd.	96359
Ap #838-2014 Quis, Rd.	92758
P.O. Box 450, 1030 Volutpat. St.	96982
P.O. Box 486, 6366 Molestie St.	95285

Showing 1 to 10 of 11 entries Previous 1 2 Next

# Crediti e riferimenti

- *Matteo Golfarelli, Stefano Rizzi - Data Warehouse - Teoria e pratica della progettazione (McGraw Hill)*
- [https://anonymousbi.wordpress.com/2013/12/15/  
pentaho-bi-server-5-0-ice-mysql-installation-guide/](https://anonymousbi.wordpress.com/2013/12/15/pentaho-bi-server-5-0-ice-mysql-installation-guide/)
- <http://infocenter.pentaho.com/help/index.jsp>
- <http://caccio.bimodeler.com/archives/337>