

# Gestione informatica dei dati

Il datawarehouse

Prof. Roberto Foderà

Cattedra di Gestione informatica dei dati  
Dipartimento di Giurisprudenza - Palermo

A.A. 2021-2022



**LUMSA**  
UNIVERSITÀ

# Il datawarehouse

---

Dall'archivio al datawarehouse:

1) il sistema informatico è al centro della capacità di prendere decisioni da parte del decision maker;

2) I database management system (DBMS) facilitano l'accesso ai dati memorizzati;

3) I data base tradizionali sono mirati a specifiche applicazioni, cioè mirano a fornire un supporto ad un processo applicativo o a specifiche funzioni: ad esempio predisporre un bilancio o analizzare le consistenze di un magazzino.

# Il datawarehouse

---

Le aziende trasferiscono regolarmente i dati archiviati dai diversi sistemi informativi, se non addirittura operativi, in questi magazzini con lo scopo di tenerli pronti per eventuali successive analisi strategiche nell'ambito della Business Intelligence (BI).

Le imprese possono disporre tendenzialmente di:

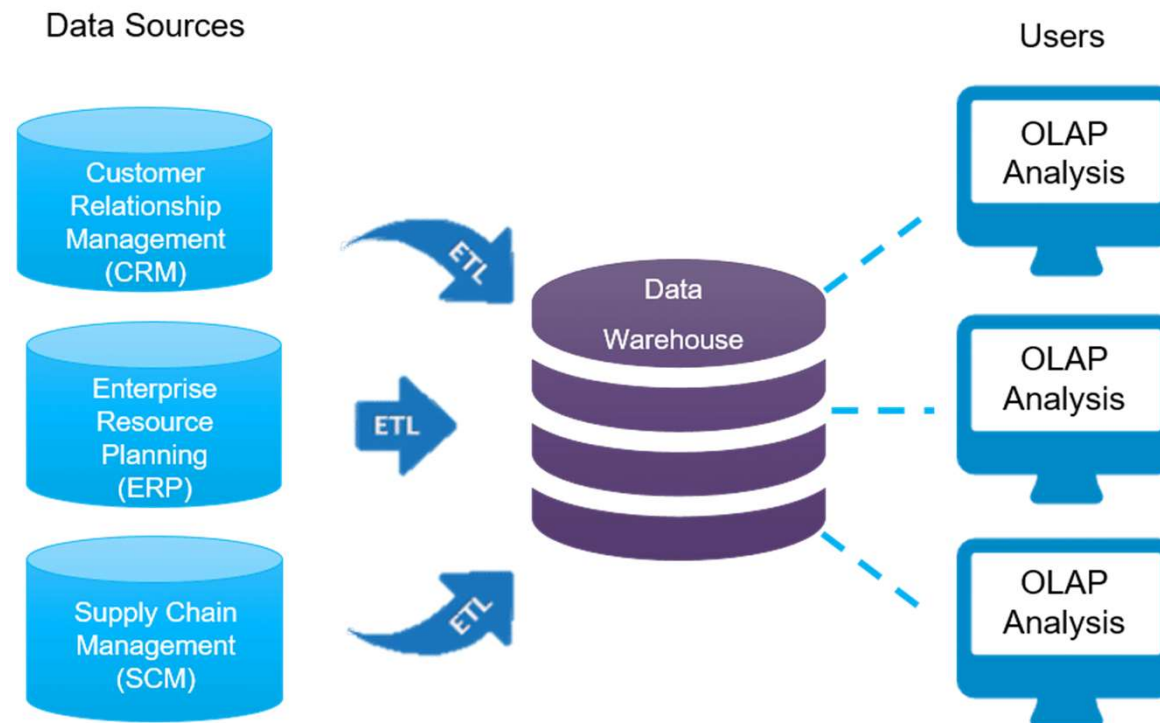
**Dati operativi:** sono informazioni transazionali generate dai sistemi amministrativi e contabili nel corso delle attività quotidiane. Derivano ad esempio dai programmi di gestione contabile o di gestione merci, o dall'Enterprise resource planning (ERP, in italiano "pianificazione delle risorse d'impresa") ecc.

**Dati pianificati:** sono i dati operativi ottenuti in seguito al processo di raccolta, archiviazione e organizzazione per l'analisi.

## Il datawarehouse

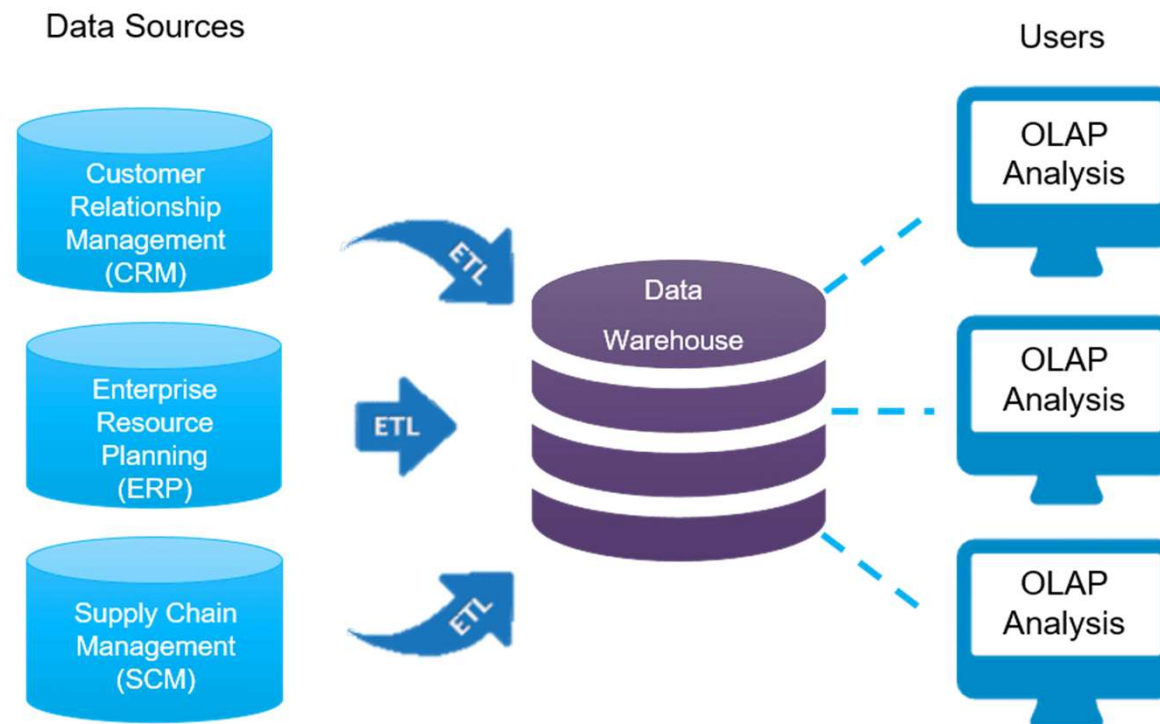
---

Un datawarehouse (DW), letteralmente in italiano magazzino di dati, permette di produrre analisi maggiormente approfondite (relazioni tra i dati) e significativamente più adeguate per il decision maker. Raccoglie, organizza e archivia dati storici provenienti da fonti diverse ed eterogenee.



## Il datawarehouse

Un DWH offre agli analisti una visione completa di dati eterogenei e consente l'aggregazione di numeri operativi chiave nel campo dell'On-Line Analytical Processing (OLAP), fungendo da punto di raccolta centrale di tutti i dati che servono al knowledge management. Un magazzino di dati funziona inoltre come base per il data mining e come riferimento per tutte le considerazioni di gestione della performance e di orientamento strategico dell'impresa.



# Il datawarehouse

---



Un datawarehouse è una collezione di dati statici integrati, organizzata per soggetti, che riguardano una serie di fatti accaduti nel tempo memorizzati in supporti non volatili e finalizzata al recupero di informazione a supporto di processi decisionali (seguendo le caratteristiche esposte da Inmon).

La costruzione di un sistema di datawarehouse non comporta l'inserimento di nuove informazioni bensì la riorganizzazione di quelle esistenti, e implica pertanto l'esistenza di un sistema informativo.

# Il datawarehouse

---



I dati sono **statici**.

In una base di dati operativa i dati vengono acceduti, inseriti, modificati, cancellati pochi alla volta.

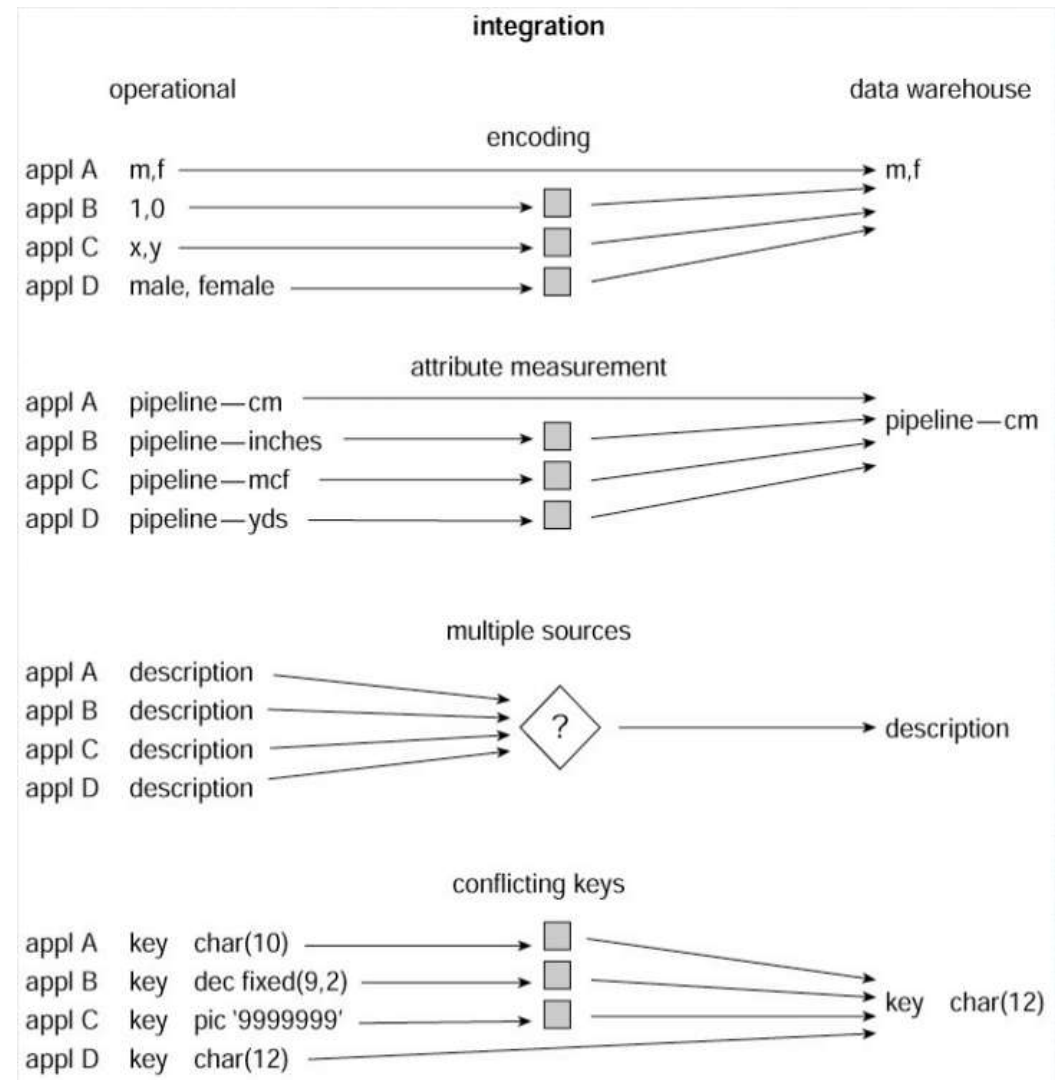
In un DW le operazioni di ricerca sono interattive mentre le operazioni di aggiornamento sono fuori linea e riguardano milioni di record.

Un DW viene considerato spesso dagli utenti come un database a sola lettura in quanto gli aggiornamenti vengono svolti “a freddo”, ossia quando il DW è fuori linea.

# Il datawarehouse

Nel DW la raccolta dati è **integrata**, in quanto in esso confluiscono dati provenienti da più fonti.

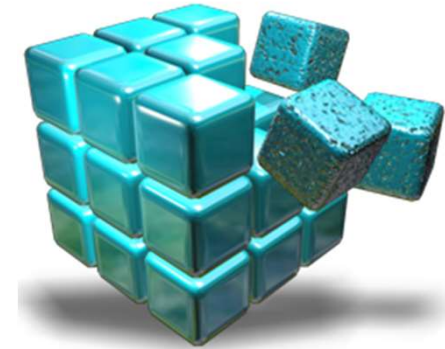
Attraverso il DW si opera una uniformità di metodi di codifica, si omogenizzano le unità di misura, si descrivono le variabili in modo congruente, ecc.





# Il datawarehouse

---



Il DW è **orientato al soggetto**.

Se, infatti, i dati in un data base sono mirati a un'applicazione, i dati in un Data Warehouse sono organizzati collegando informazioni di temi specifici (la clientela, le richieste di prodotti) per fornire un supporto ad una decisione.

La differenza deriva anche dal modo di archiviare i dati.

I database operazionali sono organizzati intorno alle differenti applicazioni del dominio aziendale; nei DW i dati vengono archiviati in modo che possano essere facilmente letti o elaborati dagli utenti cioè in modo da favorire la produzione di informazioni.

# Il datawarehouse

---



**L'orizzonte temporale** in un DW è ampio.

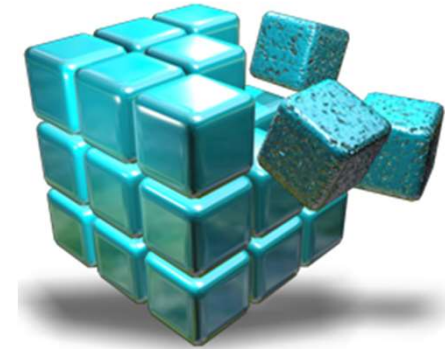
In una base di dati operativa l'orizzonte temporale di interesse è generalmente contenuto (pochi mesi, pochi anni)

In un DW l'interesse temporale è ampio: interessa l'evoluzione storica delle informazioni. Il DW deve permettere analisi che spaziano su un ampio spettro temporale.

In linea di principio non vengono mai eliminati dati dal DW.

# Il datawarehouse

---



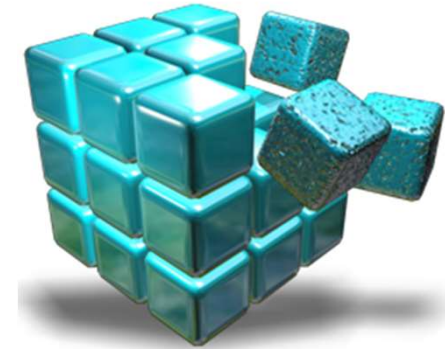
Differenze fondamentali tra database operazionali e DW si rilevano nelle tipologie di interrogazioni.

Nei data base le interrogazioni eseguono transazioni che in genere leggono e scrivono un ridotto numero di record da diverse tabelle legate da semplici relazioni: per esempio, si ricercano i dati di un cliente per inserire un suo nuovo ordine. Questo tipo di elaborazione viene comunemente detto On-Line Transactional Processing (**OLTP**).

Nei DW il tipo di elaborazione viene detto On-Line Analytical Processing (**OLAP**), ed è caratterizzato da un'analisi dinamica e multidimensionale che richiede la scansione di un'enorme quantità di record per calcolare un insieme di dati numerici di sintesi.

## Il datawarehouse

---



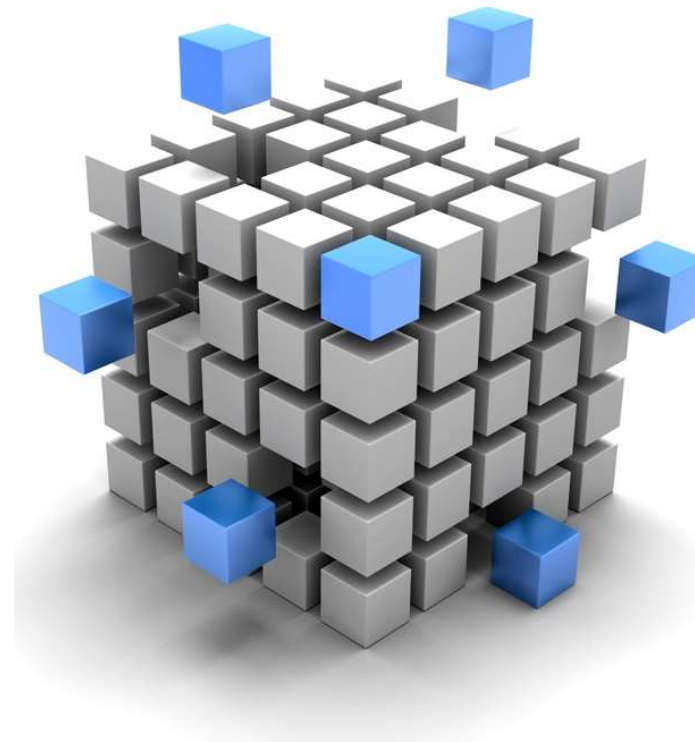
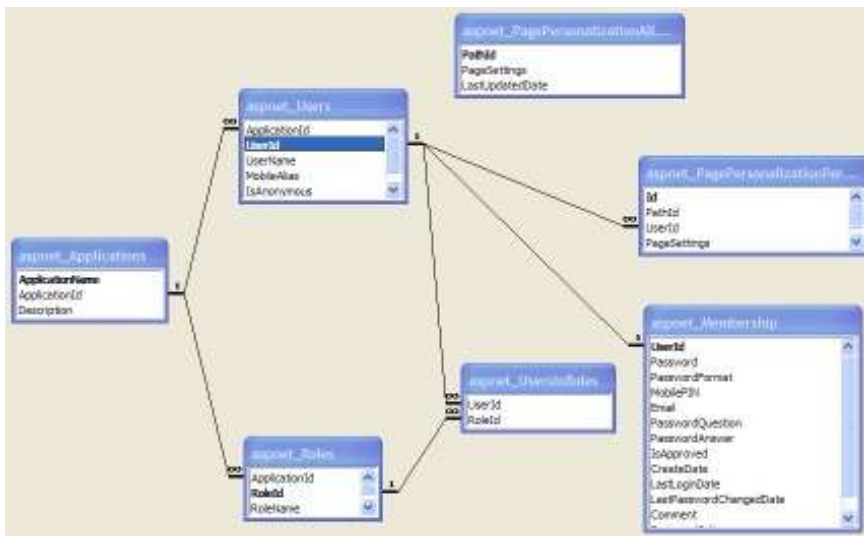
### OLAP (On-Line Analytical Processing)

Mentre gli utenti degli strumenti di reportistica svolgono un ruolo essenzialmente passivo, gli utenti OLAP sono in grado di costruire attivamente una sessione di analisi complessa in cui ciascun passo effettuato è conseguenza del passo effettuato al passo precedente.

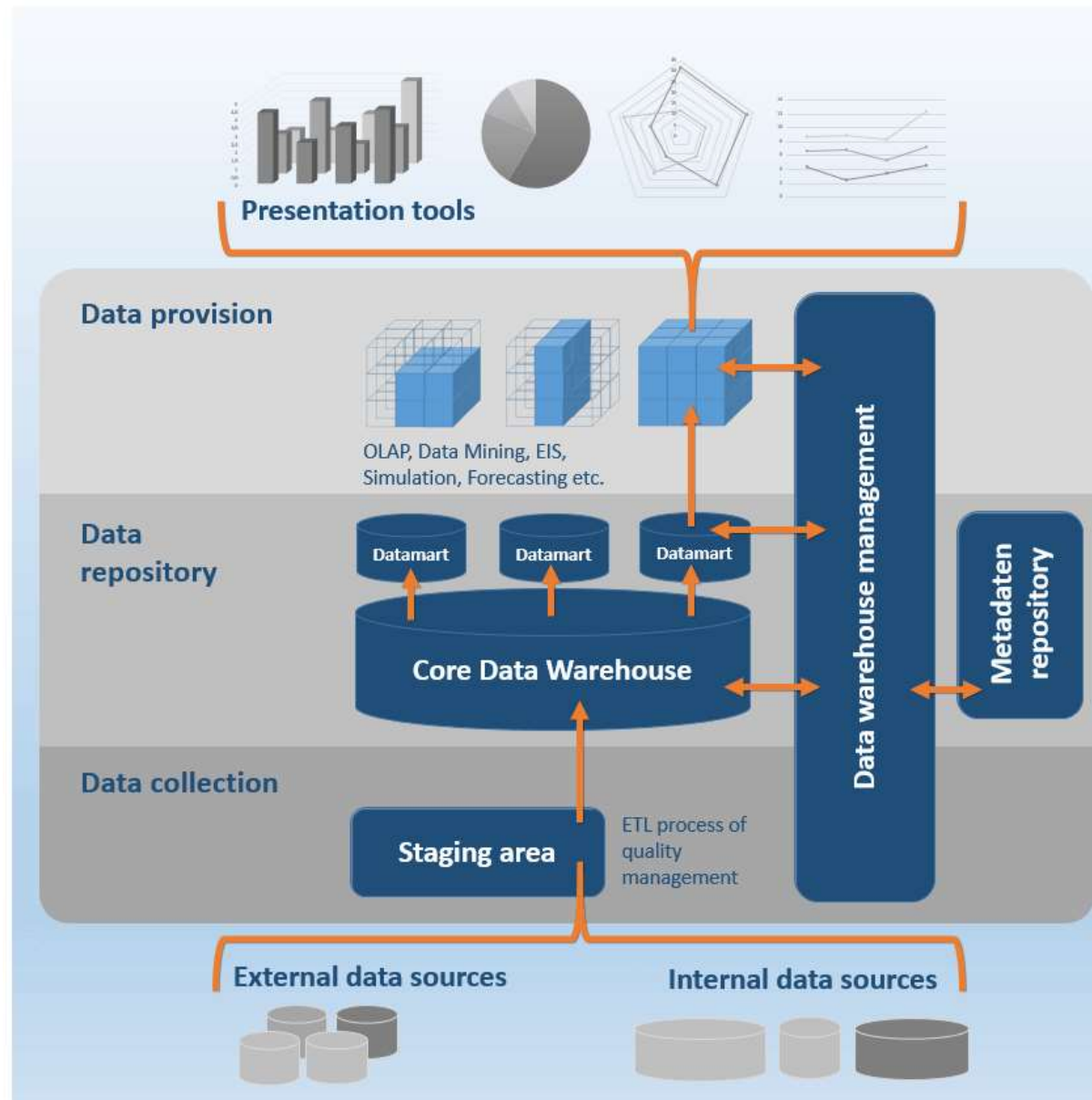
Una sessione OLAP consiste in pratica in un percorso di navigazione che si concretizza in una sequenza di interrogazioni che spesso non vengono formulate direttamente, ma per differenza rispetto a un'interrogazione precedente.

# Il datawarehouse

Una base di dati multi-dimensionale, rappresenta i dati in essa contenuti come fossero dimensioni e non tabelle.  
È più facile da usare e più veloce di una base di dati relazionale.



## Il datawarehouse



# Il datawarehouse

Il risultato delle interrogazioni è di tipo multidimensionale.

Pur descrivendo dati a più dimensioni gli strumenti OLAP rappresentano tipicamente i dati in modo tabellare, evidenziando le diverse dimensioni mediante intestazioni multiple, colori, popup di selezione, eccetera.

Popolazione residente al 1° gennaio

Personalizza Esportazioni Grafici La tua interrogazione

→ Età	totale		
→ Stato civile	totale		
→ Tipo di indicatore demografico	popolazione al 1° gennaio		
→ Anno	2016		
→ Sesso	maschi	femmine	totale
	▲ ▼	▲ ▼	▲ ▼
→ Territorio			
■ Italia	29 456 321	31 209 230	60 665 551
■ Sicilia	2 466 730	2 607 531	5 074 261
Trapani	213 368	222 397	435 765
Palermo	615 028	656 378	1 271 406
Messina	308 730	331 945	640 675
Agrigento	216 477	228 652	445 129
Caltanissetta	132 140	139 618	271 758
Enna	82 021	87 761	169 782
Catania	541 685	573 850	1 115 535
Ragusa	158 266	161 960	320 226
Siracusa	199 015	204 970	403 985

## Il datawarehouse

---

Oltre a fornire una visione dei dati in più dimensioni, questo tipo di base dati supporta la registrazione dei dati in vari livelli di aggregazione, in quanto le dimensioni sono strutturate gerarchicamente. Le **gerarchie** rappresentano i collegamenti tra la massima granularità dei dati e la possibilità di costruire cluster aggregando tipologie simili.

La dimensione aggregante si chiama «padre», quella disaggregata «figlia». Questo permette, ad esempio, di effettuare le funzioni di **drill-down** e di **roll-up** dei dati.

prodotto

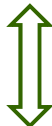


nome

anno



trimestre



mese

regione



provincia



SLL



comune





## Il datawarehouse

Una tabella presenta tipicamente due dimensioni, una esplicitata nelle intestazioni delle colonne, l'altra nelle intestazioni delle righe.

Popolazione residente al 1° gennaio 2016

		Sesso		
		maschi	femmine	totale
Territorio	Sicilia	2.466.730	2.607.531	5.074.261
	Trapani	213.368	222.397	435.765
	Palermo	615.028	656.378	1.271.406
	Messina	308.730	331.945	640.675
	Agrigento	216.477	228.652	445.129
	Caltanissetta	132.140	139.618	271.758
	Enna	82.021	87.761	169.782
	Catania	541.685	573.850	1.115.535
	Ragusa	158.266	161.960	320.226
	Siracusa	199.015	204.970	403.985

## Il datawarehouse

Se volessimo aggiungere variabili potremmo specificarle anch'esse in intestazione di righe e colonne, ampliando le celle rappresentate.

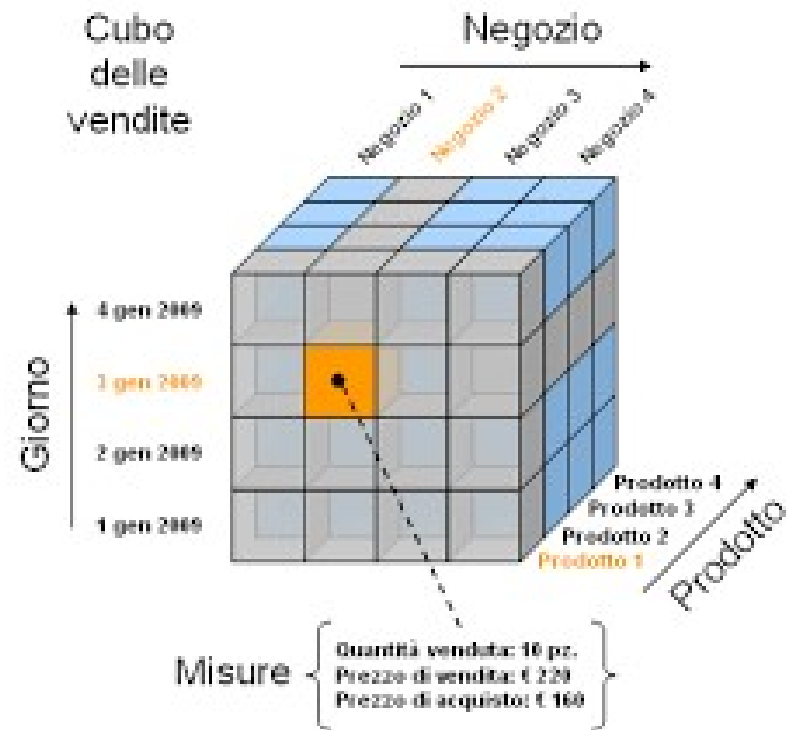
Popolazione residente al 1° gennaio 2016

		Sesso							
		maschi				femmine			
		Stato civile							
		nubile/ celibe	coniugata/ o	divorziata/ o	vedova/o	nubile/ celibe	coniugata/ o	divorziata/ o	vedova/o
Territorio	Sicilia	1.130.337	1.250.826	27.627	57.940	1.001.187	1.257.198	44.808	304.338
	Trapani	95.854	110.491	1.979	5.044	81.112	110.072	3.059	28.154
	Palermo	282.005	311.877	6.824	14.322	257.881	314.941	11.800	71.756
	Messina	140.836	155.276	4.255	8.363	127.111	156.180	6.778	41.876
	Agrigento	98.099	111.685	1.733	4.960	85.082	112.162	2.412	28.996
	Caltanissetta	59.767	67.924	1.405	3.044	52.795	67.410	2.085	17.328
	Enna	37.012	41.975	847	2.187	33.067	42.178	1.230	11.286
	Catania	253.356	270.599	5.890	11.840	227.474	273.994	10.402	61.980
	Ragusa	72.583	80.506	1.792	3.385	60.607	79.822	2.719	18.812
	Siracusa	90.825	100.493	2.902	4.795	76.058	100.439	4.323	24.150

## L'ipercubo

Si può immaginare di estendere il ragionamento ottenendo tabelle con molte variabili e, pertanto, moltissime celle.

Un ipercubo è una struttura concettuale in cui ogni «lato» è una variabile (dimensione), ogni «cubetto» è il valore corrispondente.

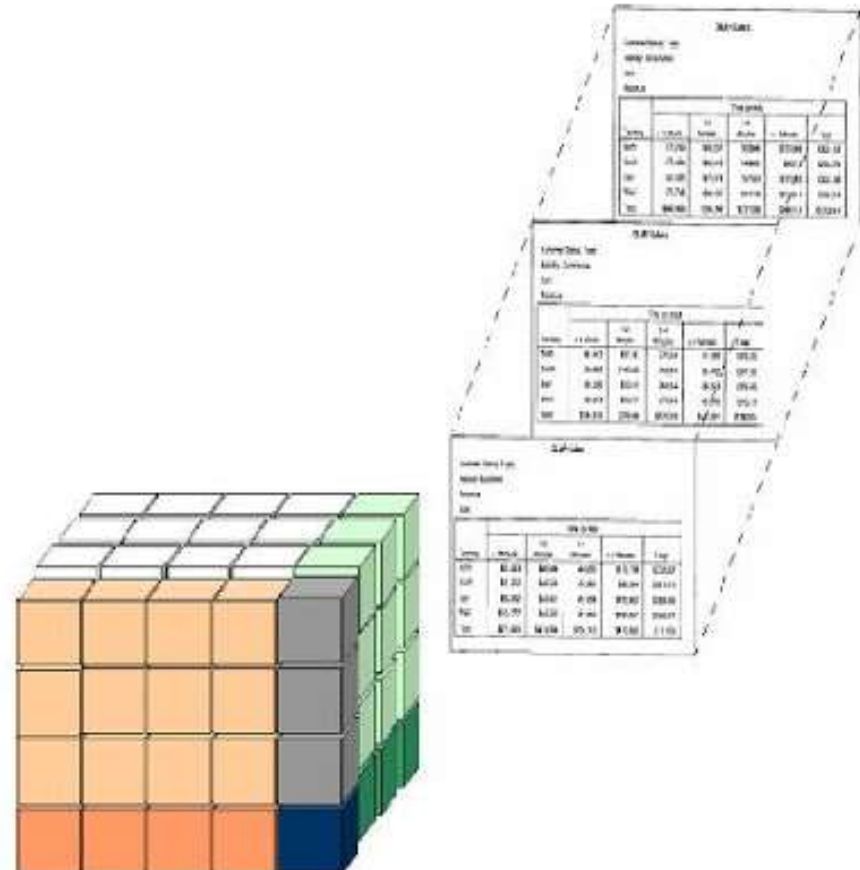


# Il datawarehouse

## L'ipercubo - dimensioni

L'ipercubo può, quindi, essere pensato come una struttura multidimensionale caratterizzata da  $n$  dimensioni. Una tabella a due variabili è un ipercubo a due dimensioni, una tabella a tre variabili è un ipercubo tridimensionale (ovvero un cubo), una tabella che considera  $n$  variabili è un ipercubo  $n$ -dimensionale.

Gli ipercubi sono il modo di rappresentare i dati in un DW.

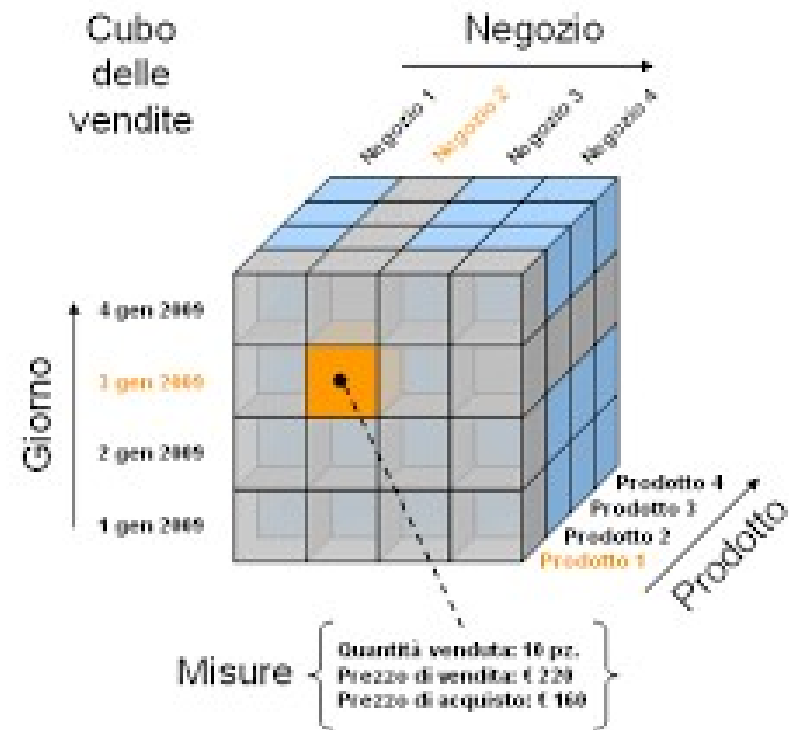


# Il datawarehouse

## L'ipercubo - dimensioni

Le dimensioni sono grandezze a valori discreti che rappresentano le prospettive di analisi dei fatti e li individuano all'interno di un opportuno contesto.

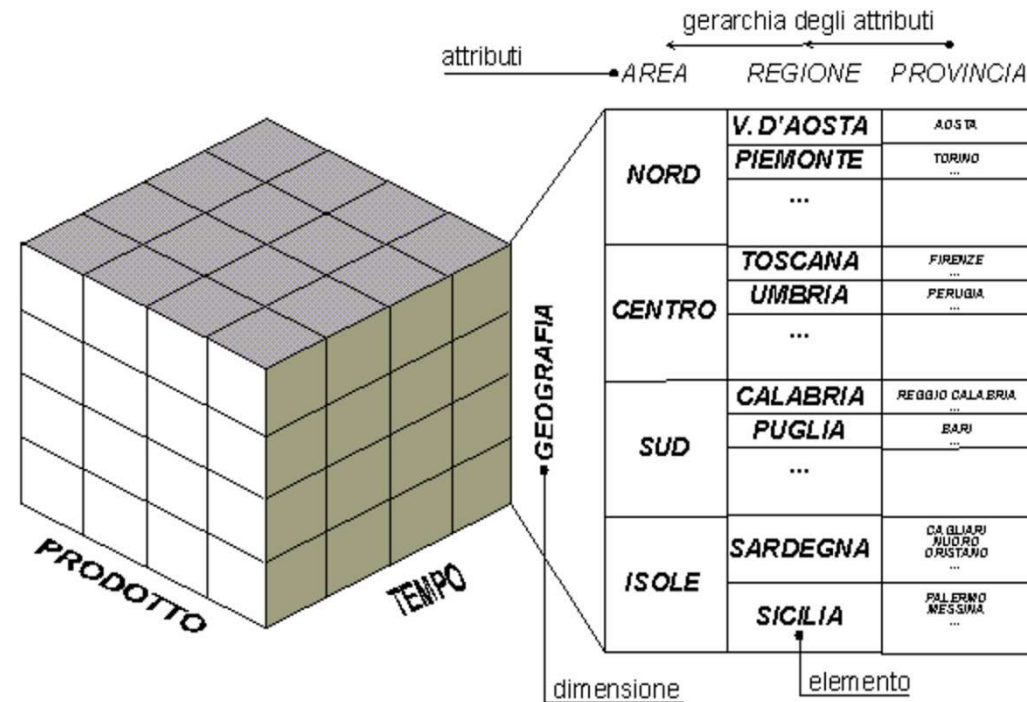
Nell'esempio mostrato, del cubo delle vendite, esse sono il giorno (tempo), il negozio e il prodotto venduto.



# Il datawarehouse

## L'ipercubo - gerarchie

Ogni dimensione può essere costituita da un insieme di attributi organizzati in opportune gerarchie, ad esempio per la collocazione dei negozi si possono gerarchizzare in provincia, regione, circoscrizione.



# Il datawarehouse

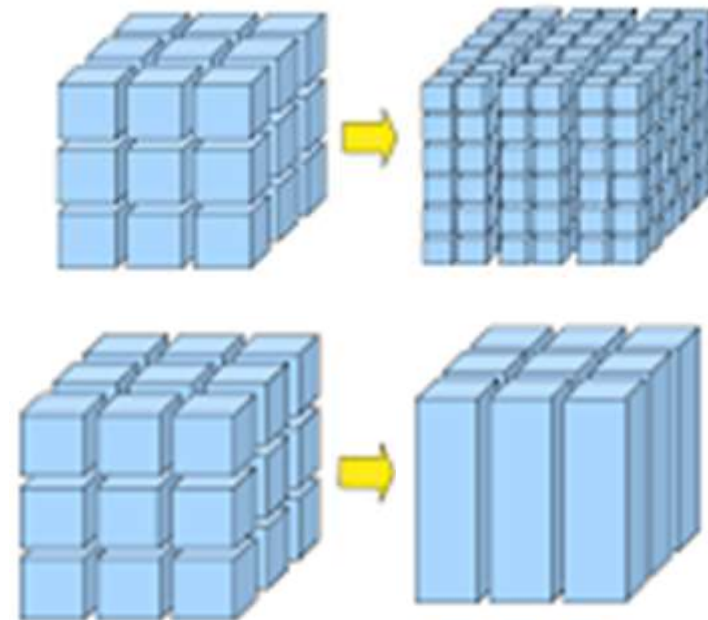
---

## L'ipercubo - gerarchie

Le gerarchie consentono di operare attraverso le funzioni di drill-down e roll-up.

Il **drill-down** permette di ampliare la richiesta dati da un livello di aggregazione «alto» a quelli dei dettagli che lo compongono.

Il **roll-up** è il procedimento inverso e «nasconde» i livelli più bassi.



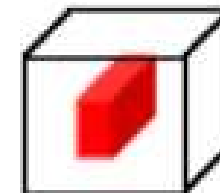
## L'ipercubo - selezione

Disporre di un ipercubo permette di selezionare dati omogenei rispetto una o più dimensioni.

**Slice:** produce “una fetta” dell’ipercubo attraverso una selezione con un vincolo su una dimensione (per esempio tutte le vendite di un tipo di prodotto presso un negozio, tutti gli indicatori demografici riferiti a una regione).



**Dice:** produce un ipercubo più piccolo estratto da quello corrente attraverso la selezione di vincoli su due o più dimensioni.

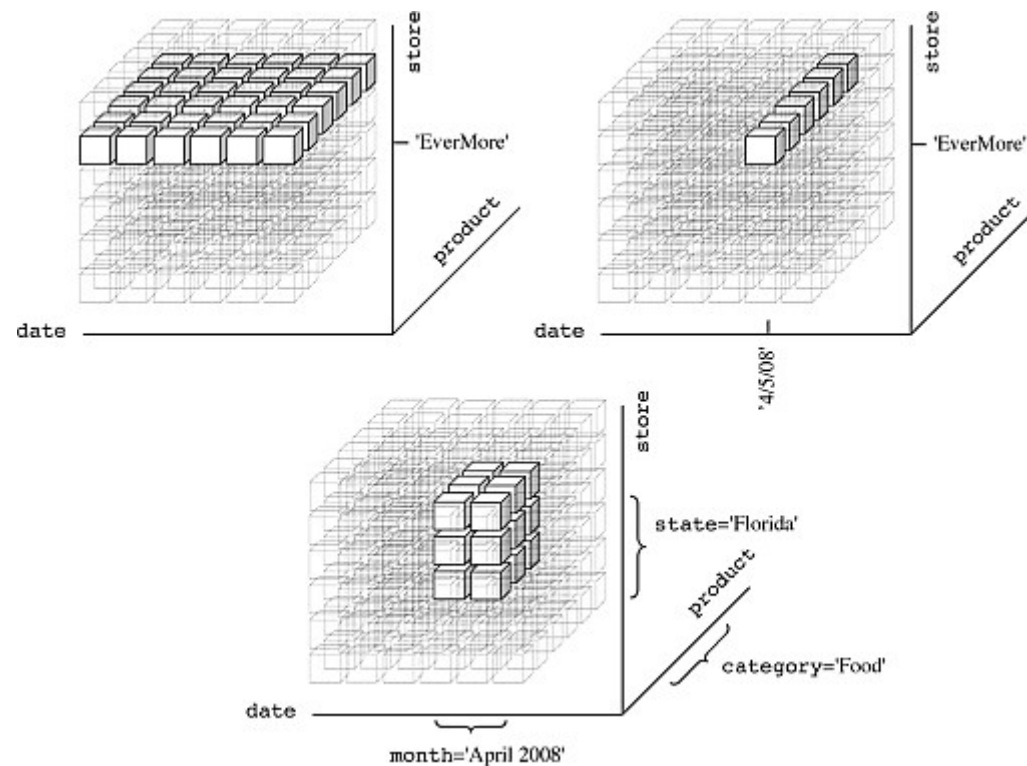




# Il datawarehouse

## L'ipercubo - selezione

Esempi di «tagli» delle dimensioni di un ipercubo.



## Componenti DW

Un datawarehouse non possiede una struttura predefinita, ma essa è determinata dai componenti che vengono utilizzati.

Un datawarehouse dispone dei seguenti componenti logici:

- Definizione dei dati (data definition)
- Acquisizione dei dati (data collection)
- Gestione dei dati (data management)
- Analisi
- Metadati

## Definizione dei dati (data definition)

Lo scopo di questo componente è la progettazione e la definizione dell'ambiente del Data Warehouse: progettare e definire la sua struttura, identificare le sorgenti di dati, definire le operazioni di pulizia dei dati e le regole di trasformazione che condizionano i dati in un formato fruibile dai processi decisionali.

Viene usato per progettare e definire la base di dati del Data Warehouse, e cioè per creare:

- Le entità (tabelle)
- Gli attributi (colonne)
- Gli identificativi (chiavi)

## Acquisizione dei dati (data collection)

Attraverso la data collection si acquisiscono i dati necessari al Data Warehouse e si predispongono i modi (piani di spoglio) e i tempi di acquisizione.

Questo componente si occupa anche della pulizia dei dati:

- Rimuovere le incongruenze
- Aggiungere dati mancanti
- Assicurarsi che l'integrità dei dati sia mantenuta

Inoltre viene effettuata anche una trasformazione dei dati:

- Aggiunta di campi relativi al tempo (es. data di estrazione)
- Aggregazione di dati di dettaglio
- Derivazione di nuovi campi

Infine i campi vengono «mappati»

## Gestione dei dati (data management)

Il data management fornisce i servizi di lavoro agli altri componenti, gestisce tutte le basi di dati all'interno del Data Warehouse.

Questo componente permette di :

- Derivare nuovi dati sommarizzati partendo da dati di dettaglio,
- Distribuire dati verso le postazioni di lavoro degli utenti,
- Applicare criteri di sicurezza,
- Svolgere operazioni di ripristino in caso di perdita dei dati
- Archiviare i dati.

La componente di gestione delle basi di dati si incarica di creare, accedere, estrarre, mantenere i dati presenti in tutto il Data Warehouse. Questo comporta il dover processare grossi volumi di dati in modo efficiente e, contemporaneamente, supportare accessi paralleli e sofisticati criteri di indicizzazione.

## Analisi

Permette l'ottenimento dei benefici derivanti dall'implementazione di un Data Warehouse. Fornisce il necessario supporto per il conseguimento dei dati organizzati da parte dell'utente e alla loro analisi.

Questo supporto consiste nel fornire accesso diretto ai dati, visualizzarli attraverso viste multidimensionali e permettere interrogazioni predefinite o determinate dall'utente stesso.

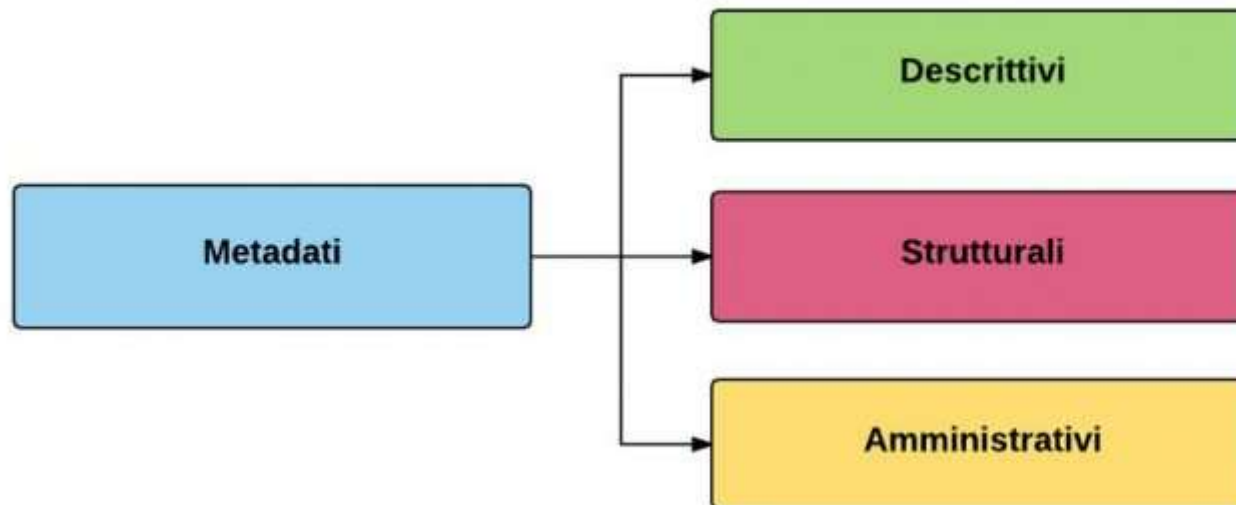
Include gli strumenti di tipo OLAP finalizzati all'analisi dei dati contenuti nel Data Warehouse. Questi strumenti sono concepiti per velocizzare l'ottenimento, la sommarizzazione, l'analisi dei dati, e per presentare una vista multidimensionale.

# Il datawarehouse

---

## Metadati

Rappresentano le informazioni che descrivono tutti i dati contenuti in un Data Warehouse. I metadati forniscono indicazioni sulla descrizione dei dati, sulla loro struttura e su dove essi sono registrati.



## I metadati strutturali

Il glossario di Eurostat, l'Ufficio di statistica dell'UE, definisce così i metadati strutturali:

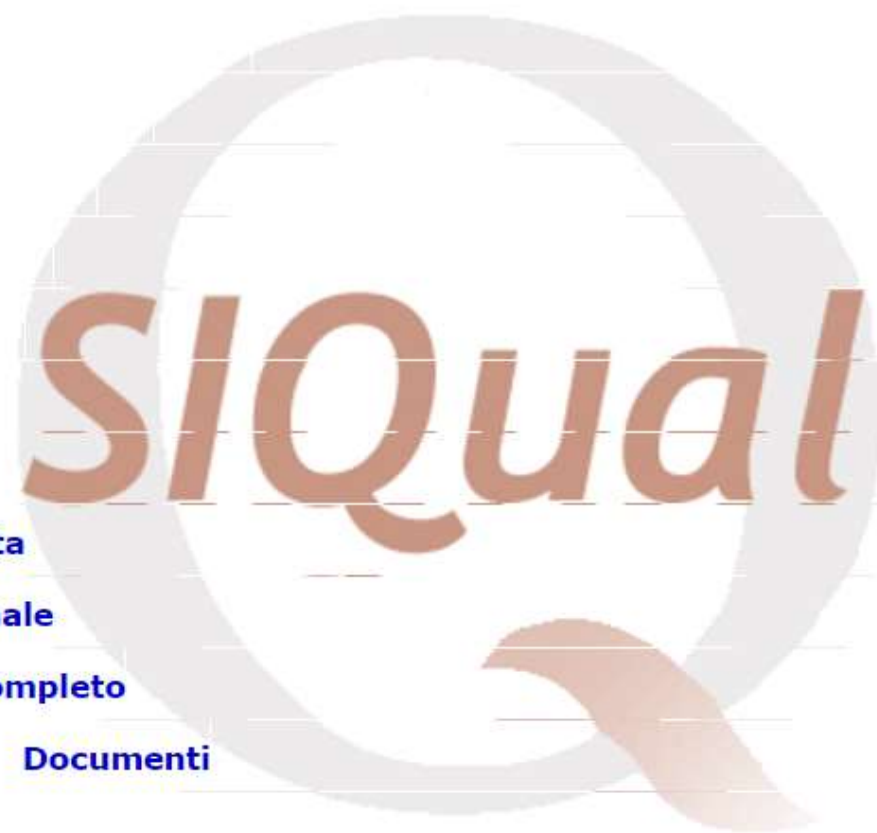
“I metadati strutturali sono utilizzati per identificare, descrivere formalmente o recuperare i dati statistici, come i nomi delle dimensioni, i nomi delle variabili, i dizionari, le descrizioni tecniche del set di dati, i luoghi del set di dati, le parole chiave per la ricerca di dati, ecc.. Per esempio, i metadati strutturali si riferiscono ai titoli delle variabili e delle dimensioni del set di dati statistici, così come le unità impiegate, liste di codici (ad esempio, per la codifica del territorio), i formati dei dati, i possibili intervalli di valori, le dimensioni temporali, gli intervalli di valori dei marcatori, le classificazioni utilizzate ecc..”



## I metadati strutturali

Il Sistema Informativo sulla Qualità (SIQual) è dedicato alla navigazione dei metadati che descrivono il processo produttivo e le sue caratteristiche: contenuto informativo, scomposizione in fasi e operazioni, attività di prevenzione, controllo e valutazione dell'errore.

**Scelta guidata**  
**Ricerca multidimensionale**  
**Elenco completo**  
**Documenti**





# I.Stat - l'accesso alle statistiche del Paese

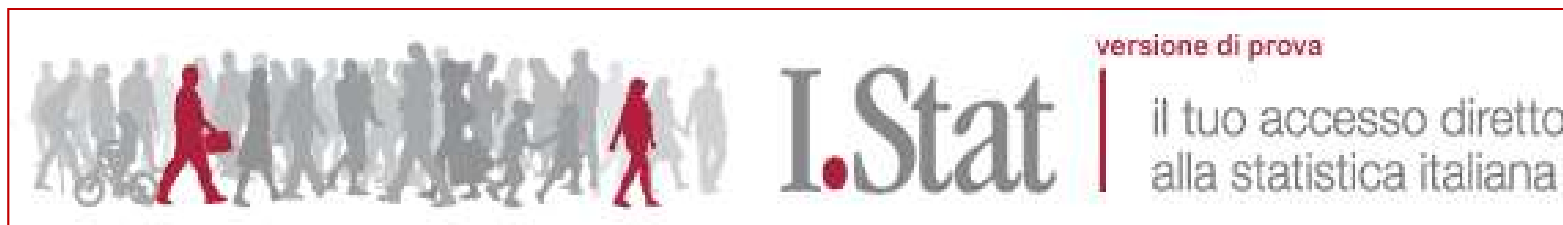
L'uso di un datawarehouse



**LUMSA**  
UNIVERSITÀ

## Il datawarehouse

---



I.Stat è il datawarehouse (letteralmente “magazzino dei dati” o in altre parole archivio informatico) delle statistiche prodotte dall’Istat. I dati sono presentati in tavole multidimensionali predefinite ma presentano un’ampia possibilità di gestione e composizione personalizzata delle tabelle

L’accesso a I.Stat è libero e gratuito per tutti all’indirizzo:  
**<http://dati.istat.it>**

Attraverso un esempio si illustreranno le principali azioni per la costruzione e l’esportazione di una tabella personalizzata.

## Obiettivo

Vogliamo visualizzare ed esportare sul nostro pc una tabella con gli indicatori statistici del tasso di conseguimento delle lauree triennali e magistrali, per tutte le regioni e sesso, relativi agli anni 2016 e 2017.

Titolo della tabella	
Intestazioni di riga	Intestazioni di colonna
	Dati
Note	
Fonte	

La tabella deve riportare:

- in fiancata, intestazione di riga, le regioni italiane e il sesso;
- in testata (intestazione di colonna), gli anni e i tassi di conseguimento delle lauree triennali e delle lauree specialistiche.

# Il datawarehouse

Sceita dell'argomento

**Esplora Temi**

Cerca nei temi  >> Annulla

Tutti i temi ▼

- Censimento agricoltura 2010
- Censimenti permanenti della popolazione e delle abitazioni, imprese, istituzioni pubbliche e nonprofit
- Censimento popolazione e abitazioni 2011
- Ambiente ed energia
- Caratteristiche del territorio
- Popolazione e famiglie
- Condizioni economiche delle famiglie e disuguaglianze
- Salute e sanità
- Assistenza e previdenza
- Istruzione e formazione
- Cultura, comunicazione, viaggi
- Giustizia e sicurezza

*Benvenuto in I.Stat, la banca dati*

**Per accedere direttamente ai dati più**

Si può consultare la banca dati [StatBase](#)

**Come si usa I.Stat**

I dati sono presentati sotto forma di tabelle. Gli

- modificarle
- esportarle
- interrogare direttamente i dati machine-

Per approfondimenti sui termini statistici utilizzi

**Altre banche dati derivanti da I.Stat**

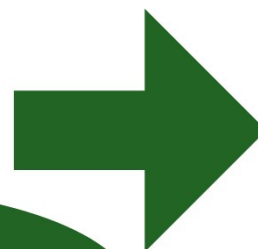
- [Giovani.Stat](#)
- [Anziani.Stat](#)
- [Immigrati.Stat](#)
- [Congiuntura.Stat](#)
- [PubblicaAmministrazione.Stat](#)
- [Capitale Immaginario.Stat](#)

Dal menù  
a sinistra si  
seleziona  
l'argomento  
che si  
intende  
esplorare

Scelta dell'argomento

Selezionando  
**Istruzione e  
formazione** si  
apre un  
ulteriore menù

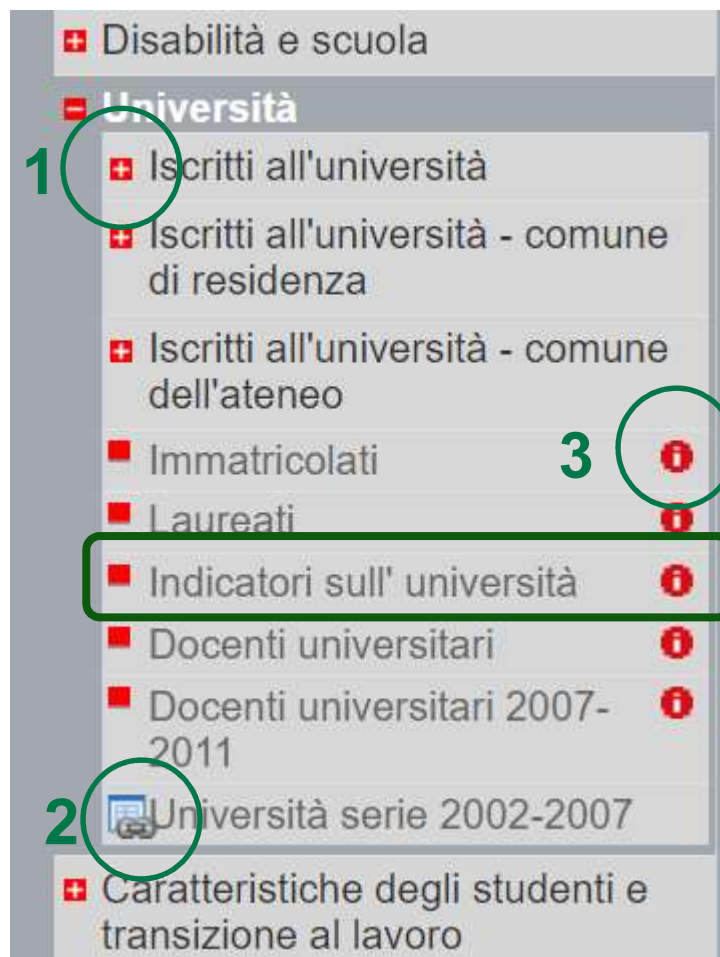
Selezionare  
**Università**



- + Ambiente ed energia
- + Caratteristiche del territorio
- + Popolazione e famiglie
- + Condizioni economiche delle famiglie e disuguaglianze
- + Salute e sanità
- + Assistenza e previdenza
- **Istruzione e formazione**
  - + Scuole
  - + Scuole (serie interrotte)
  - + Disabilità e scuola
  - + **Università**
  - + Caratteristiche degli studenti e transizione al lavoro
  - + Popolazione 15 anni e oltre per titolo di studio
  - + Giovani dai 18 ai 24 anni d'età che abbandonano prematuramente gli studi
- + Cultura, comunicazione, viaggi
- + Giustizia e sicurezza

# Il datawarehouse

Scelta dell'argomento



La tabella deve riportare:

1. È un sottotema che contiene tabelle;
2. Scarica una tabella in formato foglio elettronico;
3. Permette di visualizzare informazioni sulla tabella (metadati).

Appare un  
elenco di tabelle  
Selezionare  
**Indicatori  
sull'università**



# Il datawarehouse

## Tabella predefinita

Indicatori sull'università 

 Personalizza

 Esporta

 Grafici

 La tua interrogazione

Sesso		totale			
Seleziona anno (anno t= anno acc. t-1 / t)		2017			
Tipo dato	tasso di passaggio dalla scuola secondaria di secondo grado 	tasso di iscrizione 	tasso di conseguimento del primo titolo universitario 	tasso di conseguimento di una laurea magistrale 	
	 	 	 	 	
Regione di residenza					
 Italia	50.3	38.5	33.8	20.2	
 Nord-ovest	53.9	34.2	32.2	18.2	
Piemonte	52.4	35	30.7	18	
Valle d'Aosta / Vallée d'Aoste	(g)	34.9	30.8	17.1	
..	..	..	..	..	
Liguria	55.2	39.9	36.4	21.7	
Lombardia	54.4	39	32.9	17.9	

Le dimensioni dell'ipercubo sono quattro



# Il datawarehouse

## Tabella predefinita

Indicatori sull'università 

 Personalizza **3**

 Esporta

 Grafici

 La tua interrogazione

Sesso		totale			
Seleziona anno (anno t= anno acc. t-1 / t)		2017			
Tipo dato	tasso di passaggio dalla scuola secondaria di secondo grado	tasso di iscrizione	tasso di conseguimento del primo titolo universitario	tasso di conseguimento di una laurea magistrale	
Regione di residenza					
Italia		50.3	38.5	33.8	20.2
Nord-ovest		53.9	34.2	32.2	18.2
Piemonte		52.4	35	30.7	18
Valle d'Aosta / Vallée d'Aoste	(g)		34.9	30.8	17.1
		..			
Liguria		55.2	39.9	36.4	21.7
Lombardia		54.4	39	32.9	17.9

Le dimensioni possono essere poste a:

1. intestazione di righe;
2. Intestazione di colonne;
3. Filtro.

# Il datawarehouse

Tabella predefinita

Indicatori sull'università

Personalizza Esporta Grafica

Sesso Male

Seleziona anno (anno t / anno acc. t-1 / t-2)

Tipo dato

Regione di residenza

Italia Nord-ovest Piemonte Valle d'Aosta / Vallée (g)

tasso di passaggio dalla secondaria di secondo grado

Metadati

Indicatori sull'università

Sorgente

Fonte(i) dei dati usata (e)

**Corsi di laurea:** Dall'a.a. 2012/13 i dati sono desunti dall'Anagrafe Nazionale Studenti (ANS), a cura del Ministero dell'Università e della Ricerca (precedentemente i dati venivano raccolti attraverso una indagine effettuata dal Miur presso gli Atenei). Sono elaborati e diffusi i dati su immatricolati, iscritti e laureati per corso di laurea e gruppo disciplinare.

**Corsi di diploma universitario:** Indagine a cura del Ministero dell'Università e della Ricerca (Indagine "Definitiva"). Vengono elaborati e diffusi i dati su immatricolati, iscritti e diplomati nei corsi di diploma universitario del vecchio ordinamento. Corsi soppressi con la legge di riforma n. 509 del 1999.

**Gli indicatori relativi agli anni 2008, 2009, 2010 e 2011 sono stati ricalcolati utilizzando la popolazione statistica ricostruita su base censuaria e pertanto possono differire da quelli riportati in altre precedenti pubblicazioni.**

Corsi di laurea

Attraverso un click sulla "i" alla fine del nome della tabella, compare il riquadro dei Metadati, informativo circa i contenuti dell'insieme di dati prescelto

Modifica del formato predefinito



Attraverso il menù ***Personalizza*** è possibile modificare la tabella, selezionando le **variabili di classificazione** desiderate del fenomeno, **così come la disposizione degli assi**



## **Selezione ...**

attraverso il menù Selezione è possibile operare le scelte su tutte le dimensioni dell'ipercubo disponibile nel datawarehouse e che compone la tavola presentata.

Osserviamo le quattro dimensioni di interesse:



## **Regione di residenza [28 / 28]**

Variabile di classificazione che permette di scegliere per quale porzione di territorio visualizzare gli indicatori





## Tipo dato [4 / 4]

Variabile di classificazione che permette di l'indicatore tra quelli (4) disponibili sul datawarehouse



## Sesso [3]

Variabile di classificazione che permette di scegliere il genere delle unità di analisi (laureati)

# Il datawarehouse

Modifica del formato predefinito

The screenshot shows a web interface for a data warehouse. At the top, there are three tabs: 'Personalizza', 'Esporta', and 'Grafici'. Below these, there is a red button labeled 'Seleziona.....' which has a dropdown menu open. The dropdown menu contains several options, each preceded by a red arrow icon. The options are: 'Regione di residenza [28 / 28]', 'Tipo dato [4 / 4]', 'Sesso [3 / 3]', and 'Seleziona anno (anno t= anno acc. t-1 / t) [1]'. The 'Seleziona anno' option is highlighted with a green box, and a green arrow points from it to the 'Selezione delle date' button at the bottom right. Below the dropdown menu, there is a section titled 'Tipo dato' with a 'Ricerca' button. At the bottom, there are two radio buttons: 'Seleziona il periodo' and 'Seleziona i dati più recenti'. The 'Seleziona i dati più recenti' option is selected. Below these, there are input fields for 'Da:' and 'A:', both set to '2017'. There is also a checkbox for 'Annuale' which is checked, and a dropdown for 'Seleziona gli ultimi:' set to '1' with the unit 'anni'.

## Anno indagine [6]

Indica l'anno a cui fa riferimento il valore della statistica



Altre selezioni: **Regione di residenza [28/28]** Tipo dato[4/4] **Sesso [3/3]** **Seleziona anno (anno t= anno acc. t-1 / t) [1]** Numero di celle selezionate: 336

Trova nella selezione  ☒ Parola o frase esatta ☐

☒ Tipo dato

- ☒ tasso di passaggio dalla scuola secondaria di secondo grado
- ☒ tasso di iscrizione
- ☒ tasso di conseguimento del primo titolo universitario
- ☒ tasso di conseguimento di una laurea magistrale

Come richiesto selezioniamo i soli due indicatori (Tipo dato) che desideriamo visualizzare.  
Personalizza → Seleziona → Tipo dato

# Il datawarehouse

Riduzione del dettaglio variabili

Altre selezioni: **Regione di residenza [28/28]** Tipo dato[4/4] **Sesso [3/3]** **Seleziona anno (anno t= anno acc. t-1 / t) [1]** Numero di celle selezionate: 336

Trova nella selezione  ☒ Parola o frase esatta

☒ Tipo dato

- ☒ tasso di passaggio dalla scuola secondaria di secondo grado
- ☒ tasso di iscrizione
- ☒ tasso di conseguimento del primo titolo universitario
- ☒ tasso di conseguimento di una laurea magistrale

La selezione, quindi la visualizzazione dell'indicatore, potrà essere esclusa eliminando la spunta a sinistra.  
Escludere i primi due indicatori.

Altre selezioni: **Regione di residenza [28/28]** Tipo dato[4/4] **Sesso [3/3]** **Seleziona anno (anno t= anno acc. t-1 / t) [1]** Numero di celle selezionate: 336

Trova nella selezione   ☐ Parola o frase esatta

☐ Tipo dato

- ☒ tasso di passaggio dalla scuola secondaria di secondo grado
- ☒ tasso di iscrizione
- ☒ tasso di conseguimento del primo titolo universitario
- ☒ tasso di conseguimento di una laurea magistrale

Cliccando sul tasto **Deseleziona** si azzereranno tutte le selezioni. Ovviamente l'inverso farà il tasto Seleziona. Non è possibile non selezionare alcun valore per una dimensione.

### Selezione delle dimensioni

Altre selezioni: **Regione di residenza [28/28]** Tipo dato[4/4] **Sesso [3/3]** **Seleziona anno (anno t= anno acc. t-1 / t) [1]** Numero di celle selezionate: 336

Trova nella selezione   ☐ Parola o frase esatta

☒ Tipo dato

- ☐ tasso di passaggio dalla scuola secondaria di secondo grado
- ☐ tasso di iscrizione
- ☒ tasso di conseguimento del primo titolo universitario
- ☒ tasso di conseguimento di una laurea magistrale

Con il pulsante "Visualizza i dati" si ritorna alla tabella che apparirà modificata

# Il datawarehouse

Riduzione del dettaglio variabili

Il numero degli indicatori visualizzati è stato ridotto ai due di interesse.

Indicatori sull'università ⓘ

Personalizza ▾ Esporta ▾ Grafici ▾ La tua interrogazione ▾

Sesso	totale ▾	
Seleziona anno (anno t= anno acc. t-1 / t)	2017	
Tipo dato	tasso di conseguimento del primo titolo universitario ⓘ	tasso di conseguimento di una laurea magistrale ⓘ
	▲ ▼	▲ ▼
<b>Regione di residenza</b>		
🇮🇹 Italia	33.8	20.2
🇮🇹 Nord-ovest	32.2	18.2
Piemonte	30.7	18
Valle d'Aosta / Vallée d'Aoste	30.8	17.1
Liguria	36.4	21.7
Lombardia	32.2	17.9
🇮🇹 Nord-est	33.7	18.9
Trentino Alto Adige / Südtirol	22.9	12.8
Provincia Autonoma Bolzano / Bozen	12.7	6.7

Personalizza Esporta Grafici

**Seleziona.....**

- Disposizione degli assi
- Opzioni
- Seleziona anno**  
(anno t= anno acc. t-1 / t) [1]

Tipo dato

☐ Seleziona il periodo

Da: ☒ Annuale 2017

A: ☐ Dati più recenti 2017

☒ Seleziona i dati più recenti

Seleziona gli ultimi: 1 anni

## Seleziona anno

Per cambiare l'anno selezionare "Seleziona il periodo", indicando sotto «Da:» l'anno 2016 e sotto «A:» 2017.

# Il datawarehouse

## Indicatori sull'università

Personalizza

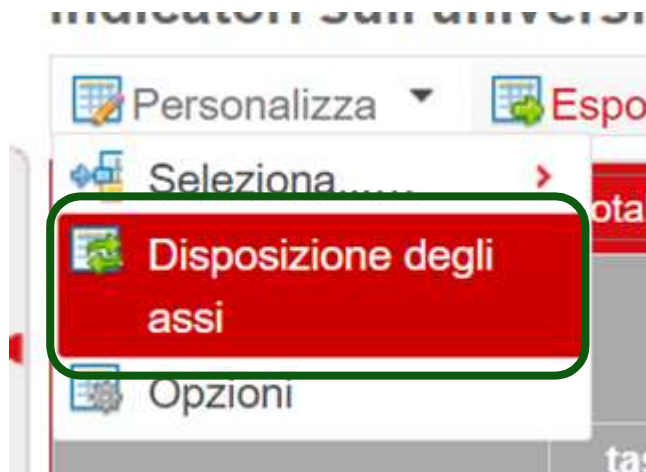
Esporta

Grafici

La tua interrogazione

Sesso	totale			
Seleziona anno (anno t= anno acc. t-1 / t)	2016		2017	
Tipo dato	tasso di conseguimento del primo titolo universitario	tasso di conseguimento di una laurea magistrale	tasso di conseguimento del primo titolo universitario	tasso di conseguimento di una laurea magistrale
Regione di residenza				
Italia	32.9	19.4	33.8	20.2
Nord-ovest	31.8	17.6	32.2	18.2
Piemonte	30.8	17.1	30.7	18
Valle d'Aosta / Vallée d'Aoste	33.7	19	30.8	17.1
Liguria	33.5	19.5	36.4	21.7
Lombardia	31.9	17.6	32.2	17.9

Abbiamo selezionato gli item che ci interessano;  
dobbiamo ora visualizzare anche la dimensione "sesso".



Per cambiare la disposizione degli assi e costruire la tabella come desiderata, selezionare Disposizione degli assi dal menù Personalizza.



# Il datawarehouse

Modifica della tabella predefinita

Dimensioni Pivot

Filtra dimensione(i)

Sesso

Dimensione(i) verticale(i):

Regione di residenza

Dimensione(i) orizzontale(i):

Seleziona anno (anno t= anno acc. t-1 / t)

Tipo dato

Opzioni

Selezione delle dimensioni

Visualizza i dati

La schermata per la disposizione degli assi propone tutte le dimensioni disponibili con, in **Dimensione(i) verticale(i)** le variabili visualizzate nelle righe, ed in **Dimensione(i) orizzontale(i)** le variabili visualizzate nelle colonne.

# Il datawarehouse

Modifica della tabella predefinita

Dimensioni Pivot

Filtra dimensione(i)

Sesso

Dimensione(i) verticale(i):

Regione di residenza

Dimensione(i) orizzontale(i):

Seleziona anno (anno t= anno acc. t-1 / t)

Tipo dato

Opzioni

Selezione delle dimensioni

Visualizza i dati

Con il tasto sinistro del mouse “prendere” la dimensione Sesso e spostarla in Dimensione(i) verticale (i).

# Il datawarehouse

Modifica della tabella predefinita

Dimensioni Pivot

Filtra dimensione(i)

Dimensione(i) verticale(i):  
Regione di residenza  
Sesso

Dimensione(i) orizzontale(i):  
Seleziona anno (anno t= anno acc. t-1 / t)  
Tipo dato

Opzioni

Selezione delle dimensioni















Visualizza i dati

Con il pulsante "Visualizza i dati" si ritorna alla tabella che apparirà modificata

# Il datawarehouse

Modifica della tabella predefinita

## Indicatori sull'università

 Personalizza  Esporta  Grafici  La tua interrogazione					
Seleziona anno (anno t= anno acc. t-1 / t)		2016		2017	
Tipo dato		tasso di conseguimento del primo titolo universitario 	tasso di conseguimento di una laurea magistrale 	tasso di conseguimento del primo titolo universitario 	tasso di conseguimento di una laurea magistrale 
					
Regione di residenza	Sesso				
 Italia	maschi	26.4	15.7	27.5	16.5
	femmine	39.6	23.3	40.3	24.1
	totale	32.9	19.4	33.8	20.2
 Nord-ovest	maschi				5.4
	femmine				1.2
	totale				8.2
Piemonte	maschi				5.4
	femmine	37.1	19.5	36.9	20.6
	totale	30.8	17.1	30.7	18
Valle d'Aosta / Vallée d'Aoste	maschi	26.6	13.8	23.7	14.8
	femmine	41.7	24.8	37.6	19.3
	totale	33.7	19	30.8	17.1

La tabella non è ancora quella desiderata poiché appaiono le ripartizioni geografiche assieme alle regioni, che sono i soli territori richiesti.

# Il datawarehouse

## Selezione del territorio

Regione di residenza

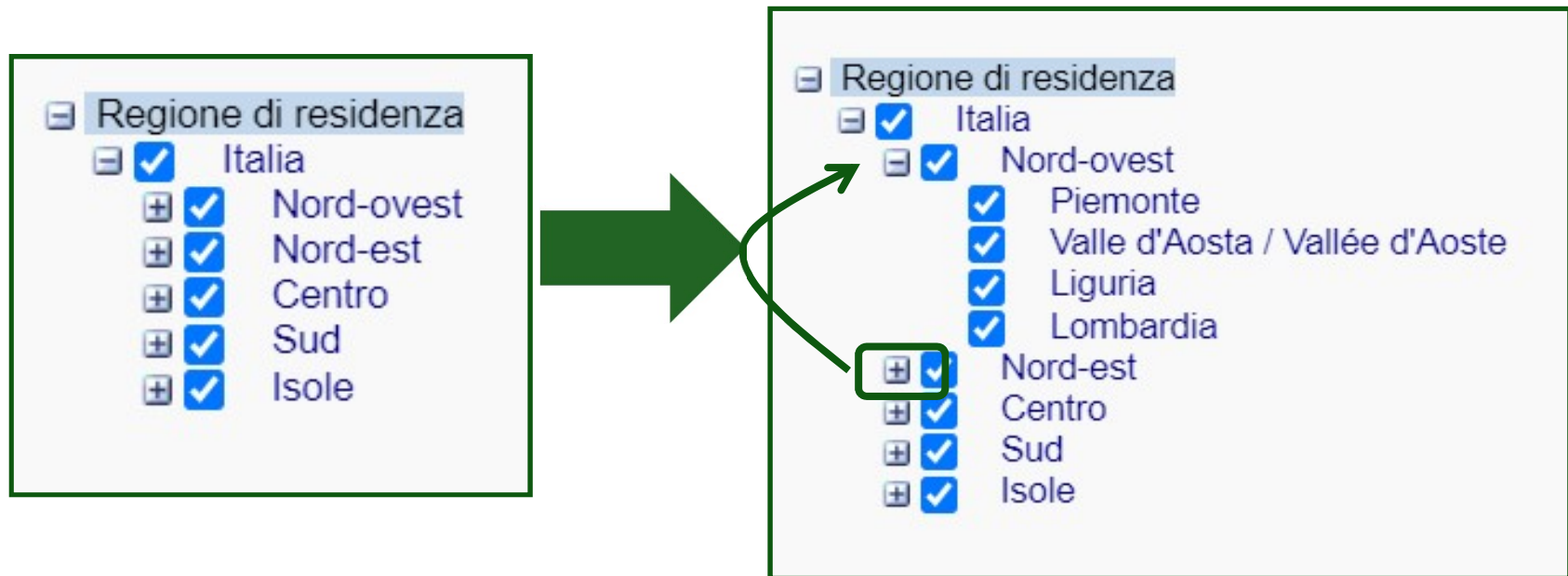
- ☒ Italia
  - ☒ Nord-ovest
    - ☒ Piemonte
    - ☒ Valle d'Aosta / Vallée d'Aoste
    - ☒ Liguria
    - ☒ Lombardia
  - ☒ Nord-est
    - ☒ Trentino Alto Adige / Südtirol
    - ☒ Provincia Autonoma Bolzano / Bozen
    - ☒ Provincia Autonoma Trento
    - ☒ Veneto
    - ☒ Friuli-Venezia Giulia
    - ☒ Emilia-Romagna
  - ☒ Centro
    - ☒ Toscana
    - ☒ Umbria
    - ☒ Marche
    - ☒ Lazio
  - ☒ Sud
    - ☒ Abruzzo

Alcune note: 1) per scorrere le regioni usare la barra di scorrimento a destra

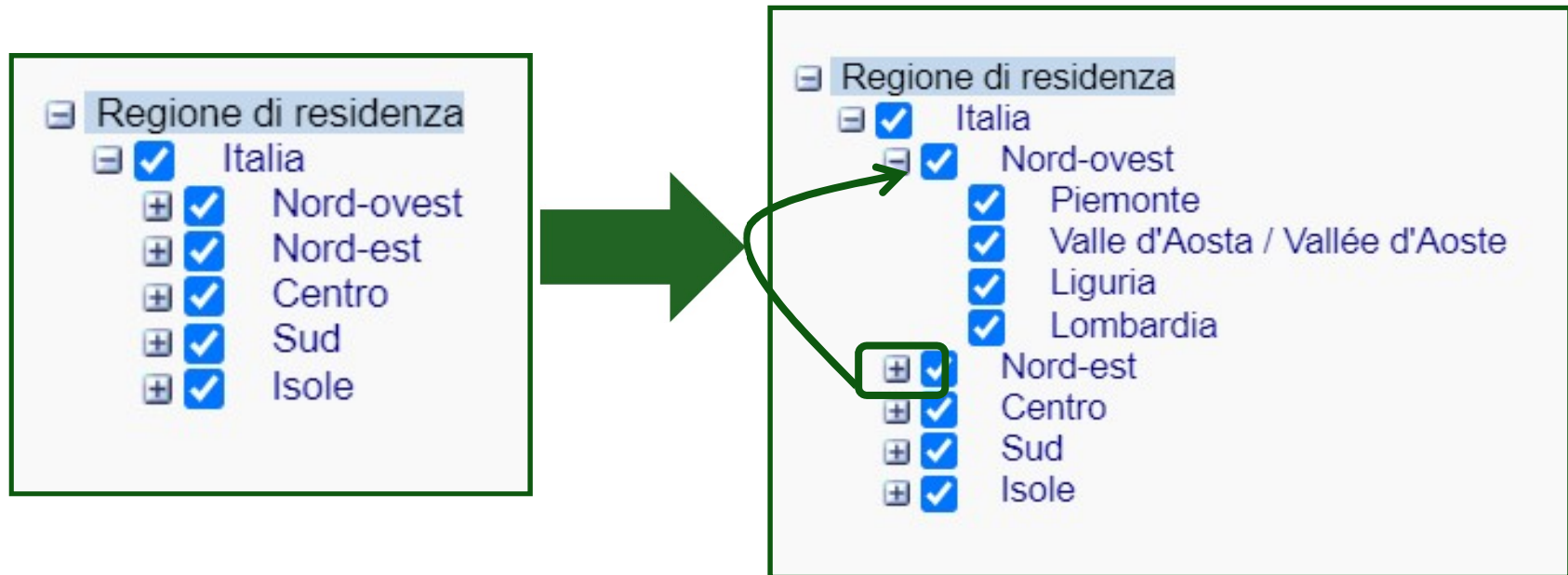
Disposizione degli assi   Opzioni   **Visualizza i dati**

Dal menù Personalizza → Selezione scegliere Regione di residenza. Levare la spunta dalle voci relative all'Italia e alle ripartizioni territoriali.

## Selezione del territorio



2) il bottone con il segno più o il segno meno accanto una voce indica la navigazione su elementi superiori o inferiori.



3) se si clicca due volte su un elemento padre (ad esempio Nord-ovest) le voci figlie (ad esempio le quattro regioni del Nord-ovest) assumeranno la stessa selezione del primo.

Regione di residenza


- ☒ Italia
  - ☐ Nord-ovest
    - ☒ Piemonte
    - ☒ Valle d'Aosta / Vallée d'Aoste
    - ☒ Liguria
    - ☒ Lombardia
  - ☐ Nord-est
    - ☒ Trentino Alto Adige / Südtirol
    - ☒ Provincia Autonoma Bolzano / Bozen
    - ☒ Provincia Autonoma Trento
    - ☒ Veneto
    - ☒ Friuli-Venezia Giulia
    - ☒ Emilia-Romagna
  - ☐ Centro
    - ☒ Toscana
    - ☒ Umbria
    - ☒ Marche
    - ☒ Lazio
  - ☐ Sud
    - ☒ Abruzzo





Dopo aver deselezionato le ripartizioni e l'Italia, con il pulsante "Visualizza i dati" si ritorna alla tabella che apparirà modificata





# Il datawarehouse

Esportazione in un foglio di calcolo

Indicatori sull'università 

Personalizza  **Esporta**  Grafici  La tua interrogazione 

Seleziona anno (anno t=)  2016 2017

anno 

Regione di residenza

 Italia

Piemonte

Valle d'Aosta / Vallée d'Aoste

maschi

femmine

totale



Excel

File di testo (CSV)

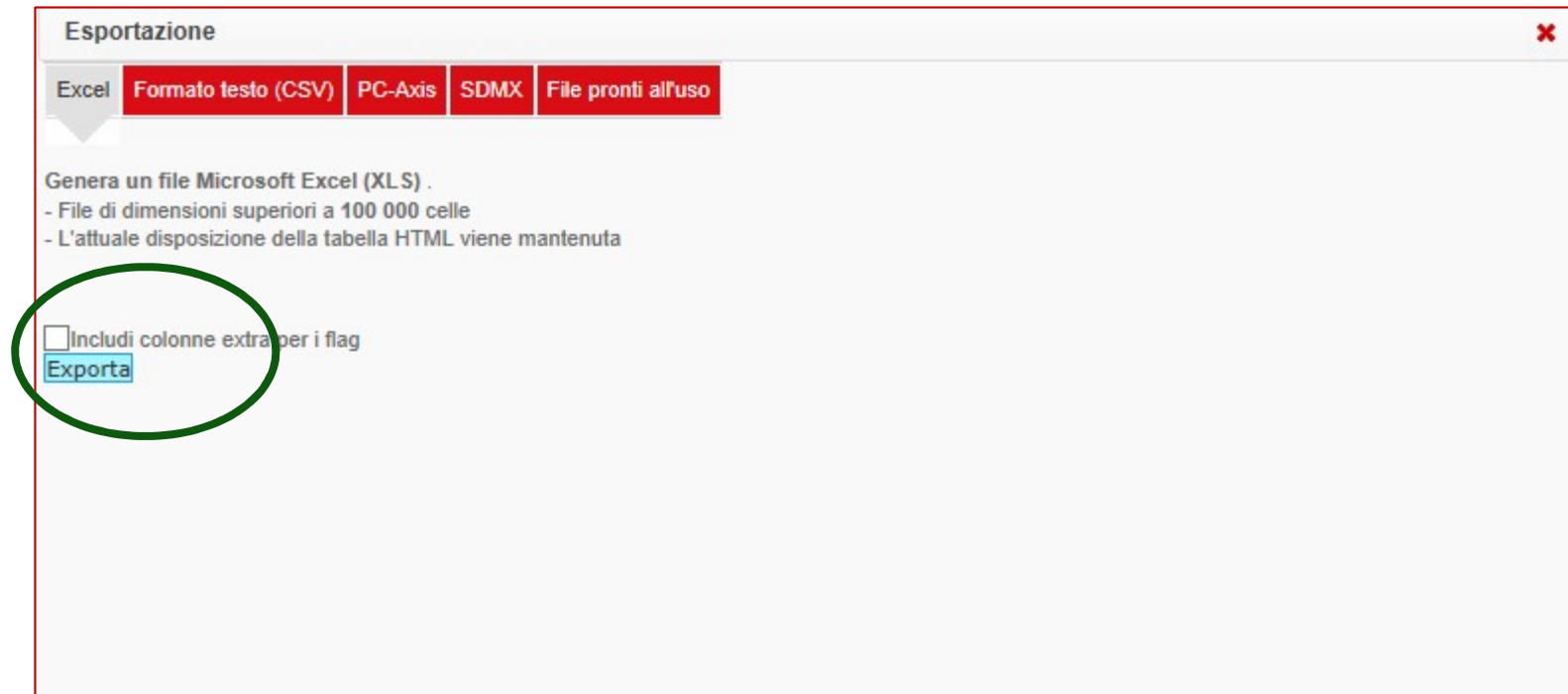
PC-axis

SDMX

File già pronti

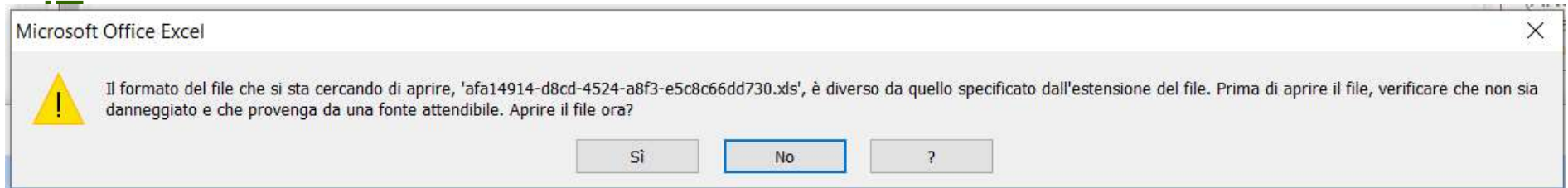
		2016	2017	
	seguimento a magistrale		tasso di conseguimento del primo titolo universitario 	tasso di conseguimento di una laurea magistrale 
		15.7	27.5	16.5
		23.3	40.3	24.1
		19.4	33.8	20.2
		14.7	24.9	15.4
		19.5	36.9	20.6
	totale	30.8	17.1	30.7
				18
	maschi	26.6	13.8	23.7
	femmine	41.7	24.8	37.6
	totale	33.7	19	30.8
				17.1

Una volta ottenuta la tabella desiderata è possibile esportarla in diversi formati (xls, csv, PC-Axis, SDMX). Per esportarla in formato xls, selezionare dal *menù* **Esporta** la voce "Excel"



Nella finestra di dialogo che si apre, selezionare **Esporta**

La tabella viene salvata in formato xml pertanto darà un messaggio di warning all'apertura con il software Excel



Con il formato xml la tabella "porta con sé" link multimediali e formati delle celle.

# Il datawarehouse

## Esportazione in un foglio di calcolo

Appunti		Carattere		Allineamento		Numeri	
A1		=DotStatQuery(B1)					
A		B	C	D	E	F	G
Dataset:Indicatori sull'università							
Seleziona anno (anno t= anno acc. t-1 / t)		2016		2017			
Tipo dato		tasso di conseguim ento del primo titolo universitari o	tasso di conseguim ento di una laurea magistrale	tasso di conseguim ento del primo titolo universitari o	tasso di conseguim ento di una laurea magistrale		
Regione di residenza		Sesso					
Italia	maschi	26,4	15,7	27,5	16,5		
	femmine	39,6	23,3	40,3	24,1		
	totale	32,9	19,4	33,8	20,2		
Piemonte	maschi	24,7	14,7	24,9	15,4		
	femmine	37,1	19,5	36,9	20,6		
	totale	30,8	17,1	30,7	18		
Valle d'Aosta / Vallée d'Aoste	maschi	26,6	13,8	23,7	14,8		
	femmine	41,7	24,8	37,6	19,3		
	totale	33,7	19	30,8	17,1		
Liguria	maschi	27,3	16,2	29,9	18,5		
	femmine	39,8	22,7	43,1	25		
	totale	33,5	19,5	36,4	21,7		
Lombardia	maschi	26,1	14,8	27	15		
	femmine	37,9	20,5	37,7	20,9		
	totale	31,9	17,6	32,2	17,9		
Trentino Alto Adige / Südtirol	maschi	18,7	11,9	16,7	10,5		
	femmine	30,9	16,5	29,4	15,3		
	totale	24,7	14,2	22,9	12,8		
Provincia Autonoma Bolzano / Bozen	maschi	9,3	6,2	8	5,4		
	femmine	17,8	6,6	17,7	8,1		

Per includere oggetti grafici o formati speciali di Excel, bisognerà salvare il file in formato cartella xlsx.