

00. Un semplice esempio di database

Il problema della ridondanza dei dati

database

un semplice esempio di database

Una **Base di Dati** (detta anche **database** o più semplicemente **DB**), non è altro che un insieme organizzato di dati in qualche modo correlati tra loro e realizzato con uno scopo ben preciso, comunque connesso con l'elaborazione dei dati al fine di ottenere informazioni; più dettagliatamente un database ha le seguenti caratteristiche:

- rappresenta un determinato aspetto della realtà osservata (ad es. un'azienda, un'organizzazione, una situazione reale...); ogni cambiamento di tale realtà si riflette quindi sulla base di dati in questione;
- è una raccolta di dati logicamente coerenti e con un significato intrinseco, quindi un assortimento casuale di dati (come l'insieme delle parole di queste slides) non può essere considerato una base di dati;
- è progettata e realizzata con uno scopo ben definito.

un semplice esempio di database

Supponiamo di voler realizzare un database per la registrazione degli esami svolti dagli studenti di un certo ateneo universitario e di volere gestire anche le informazioni riguardanti i docenti e le diverse materie di insegnamento.

Trattandosi sostanzialmente di due differenti tipologie di entità (cioè gli esami svolti dagli studenti ed i docenti afferenti ai diversi insegnamenti) potremmo pensare di organizzare due tabelle contenenti i dati richiesti.

Una tabella “REGISTRO_ESAMI” per i dati relativi all’esame sostenuto da ciascuno studente, più una tabella “INSEGNAMENTI” per registrare i docenti dell’ateneo e gli insegnamenti corrispondenti.

un semplice esempio di database

Nello specifico, la tabella “REGISTRO_ESAMI” potrebbe registrare i dati anagrafici degli studenti (numero di matricola, nome, cognome, residenza, data di nascita, ecc.), i dati relativi alla materia di insegnamento oggetto dell’esame sostenuto (un codice della materia, la denominazione, ecc.) e i dati relativi all’esame sostenuto dallo studente (la data dell’esame, il voto, le informazioni sul docente, ecc.):

Matricola	Cognome	Nome	Residenza	DataNascita	CodMateria	Denominazione	DataEsame	Esito	CodDocente	ecc.
8473	Baldieri	Marco	Catanzaro	12/07/1980	A 01	Chimica Gen.	09/04/2004	18	KR 003	...
8841	Ragusa	Ugo	Crotone	04/05/1980	A 01	Chimica Gen.	14/04/2004	27	KR 003	...
8473	Baldieri	Marco	Catanzaro	12/07/1980	B 01	Anatomia	01/06/2004	21	FT 002	...
8473	Baldieri	Marco	Catanzaro	12/07/1980	C 09	Informatica	12/06/2004	30	DM 001	...
7639	Criniti	Anna	Soverato	02/09/1975	B 01	Anatomia	23/06/2004	20	FT 002	...
8841	Ragusa	Ugo	Crotone	04/05/1980	D 02	Fisica I	11/07/2004	25	DM 001	...
...
REGISTRO_ESAMI										

un semplice esempio di database

L'altra tabella "INSEGNAMENTI" registra invece i dati anagrafici del docente (un codice identificativo, nome, cognome, data di nascita, data di assunzione, ecc.), e i dati relativi alla materia insegnata (il codice della materia, la denominazione, il monte ore, ecc.):

CodDocente	CodMateria	Denominazione	MonteOre	Cognome	Nome	DataNascita	DataAssunz.	ecc.
DM 001	C 09	Informatica	70	De Giorgis	Giovanni	13/07/1960	03/06/1998	...
DM 001	D 02	Fisica I	55	De Giorgis	Giovanni	13/07/1960	03/06/1998	...
FT 002	B 01	Anatomia	80	Perli	Piero	02/12/1955	12/03/1987	...
KR 003	A 01	Chimica Gen.	75	Tripolino	Alfredo	07/05/1962	04/09/1990	...
...
INSEGNAMENTI								

un semplice esempio di database

Volendo ad es. stabilire quali studenti abbiano superato l'esame di "Chimica Generale" con un voto superiore a 25/30, basterebbe consultare la tabella REGISTRO_ESAMI dove abbiamo a disposizione tutte le informazioni che ci interessano (comprese quelle relative al docente col quale è stato sostenuto l'esame in questione).

[illegible]

un semplice esempio di database

In realtà, lo schema proposto non è affatto efficiente, dal momento che come si può notare, ogniqualevolta uno studente supera un esame, i suoi dati anagrafici devono essere riscritti in una nuova riga della tabella REGISTRO_ESAMI; allo stesso modo, quando ad un docente viene assegnato un nuovo corso da svolgere, i suoi dati vanno riscritti, per evitare registrazioni incomplete.

Tutto ciò comporta una notevole *ridondanza* incontrollata nel database, e la ridondanza introduce ulteriori problemi di:

- *Inconsistenza Potenziale;*
- *Anomalie di Inserimento;*
- *Anomalie di Cancellazione.*

un semplice esempio di database

→ **Inconsistenza Potenziale** (*Anomalia di Aggiornamento*): come conseguenza diretta della ridondanza dei dati, nel caso ad es. cambi la città di residenza di uno studente, questa andrebbe modificata anche in tutte le righe della tabella REGISTRO_ESAMI in cui è presente lo studente, col rischio che il dato in questione non venga modificato in tutte le sue occorrenze, o che venga modificato in maniera differente in alcune altre; **Anomalia di aggiornamento**: non sappiamo se Ugo Ragusa

Anomalia di aggiornamento: non sappiamo se Ugo Ragusa è residente in “Cosenza” o “Crotone”.

Matricola	Cognome	Nome	Residenza	DataNascita	CodMateria	Denominazione	DataEsame	Esito	CodDocente	ecc.
8473	Baldieri	Marco	Catanzaro	12/07/1980	A 01	Chimica Gen.	09/04/2004	18	KR 003	...
8841	Ragusa	Ugo	Cosenza	04/05/1980	A 01	Chimica Gen.	14/04/2004	27	KR 003	...
8473	Baldieri	Marco	Catanzaro	12/07/1980	B 01	Anatomia	01/06/2004	21	FT 002	...
8473	Baldieri	Marco	Catanzaro	12/07/1980	C 09	Informatica	12/06/2004	30	DM 001	...
7639	Criniti	Anna	Soverato	02/09/1975	B 01	Anatomia	23/06/2004	20	FT 002	...
8841	Ragusa	Ugo	Crotone	04/05/1980	D 02	Fisica I	11/07/2004	25	DM 001	...
...
REGISTRO_ESAMI										

un semplice esempio di database

- **Anomalie di Inserimento:** non possiamo registrare i dati anagrafici di uno studente se questo non ha superato almeno un esame, a meno di non lasciare vuote le caselle (*campi*) relative all'esame;

Anomalia di inserimento: le caselle relative all'esame sostenuto sono vuote.



Matricola	Cognome	Nome	Residenza	DataNascita	CodMateria	Denominazione	DataEsame	Esito	CodDocente	ecc.
8473	Baldieri	Marco	Catanzaro	12/07/1980	A 01	Chimica Gen.	09/04/2004	18	KR 003	...
8841	Ragusa	Ugo	Crotone	04/05/1980	A 01	Chimica Gen.	14/04/2004	27	KR 003	...
8473	Baldieri	Marco	Catanzaro	12/07/1980	B 01	Anatomia	01/06/2004	21	FT 002	...
8473	Baldieri	Marco	Catanzaro	12/07/1980	C 09	Informatica	12/06/2004	30	DM 001	...
7639	Criniti	Anna	Soverato	02/09/1975	B 01	Anatomia	23/06/2004	20	FT 002	...
8841	Ragusa	Ugo	Crotone	04/05/1980	D 02	Fisica I	11/07/2004	25	DM 001	...
8892	Raffaelli	Pietro	Soverato	04/01/1965

REGISTRO_ESAMI

un semplice esempio di database

→ **Anomalie di Cancellazione:** se dovessimo decidere di eliminare dalla proposta formativa alcuni insegnamenti, dovremmo cancellare tutte le righe della tabella INSEGNAMENTI corrispondenti a tali corsi, con la conseguente possibile perdita dei dati anagrafici dei docenti che li svolgono.

Anomalia di cancellazione: se dovessimo eliminare gli insegnamenti di “Informatica” e “Fisica 1”, perderemmo anche i dati anagrafici del docente “De Giorgis” che insegna proprio quelle materie.

CodDocente	CodMateria	Denominazione	MonteOre	Cognome	Nome	DataNascita	DataAssunz.	ecc.
DM 001	C 09	Informatica	70	De Giorgis	Giovanni	13/07/1960	03/06/1998	...
DM 001	D 02	Fisica I	55	De Giorgis	Giovanni	13/07/1960	03/06/1998	...
FT 002	B 01	Anatomia	80	Perli	Piero	02/12/1955	12/03/1987	...
KR 003	A 01	Chimica Gen.	75	Tripolino	Alfredo	07/05/1962	04/09/1990	...
...
INSEGNAMENTI								

un semplice esempio di database

La soluzione in questi casi consiste proprio nel rendere il sistema il più irridondante possibile, creando una tabella per ogni tipo di oggetto (entità) di interesse per il lavoro che dobbiamo realizzare.

Le anomalie sono infatti quasi sempre generate dalla presenza di tipologie di oggetti di natura eterogenea nella stessa tabella.

NB: Vale la pena di precisare che in questo contesto si parlerà di “tipologia di oggetto” o di “entità” con la stessa accezione data al termine “classe” relativamente alla programmazione OOP, cioè insieme di oggetti aventi le stesse caratteristiche peculiari.

un semplice esempio di database

Possiamo dunque riorganizzare la base di dati su cinque distinte tabelle, in modo che ciascuna di esse sia **preposta alla raccolta dei dati di un solo determinato tipo di oggetto di interesse**.

Ogni tabella dunque contiene solo dati omogenei (quindi, ad es. la tabella STUDENTE contiene solo i dati di ciascuno studente, mentre la tabella ESAME i dati di ciascun esame svolto dagli studenti, ecc.).

La tabella STUDENTE registra solo i dati anagrafici degli studenti iscritti nel nostro ateneo universitario.



Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1980	02/09/2003	Catanzaro	Scienze Infermierist.	1
7639	Criniti	Anna	02/09/1975	12/10/2002	Soverato	Fisioterapista	2
8841	Ragusa	Ugo	04/05/1980	04/10/2003	Crotone	Tecn. Prevenzione	1
...
STUDENTE							

un semplice esempio di database

Matricola	CodMateria	DataEsame	Esito
8473	A01	09/04/2004	18
8841	A01	14/04/2004	27
8473	B01	01/06/2004	21
8473	C09	12/06/2004	30
7639	A01	23/06/2004	20
8841	D02	11/07/2004	25
...
ESAME			

La tabella ESAME registra i dati relativi all'esame svolto dallo studente con un determinato docente.

Da notare che dello studente viene riportata solo la matricola (che identifica in modo univoco ogni studente), così come della materia viene riportato solo il codice identificativo.

Codice	Denominazione	AnnoCorso	CorsoLaurea	Semestre
A01	Chimica Generale	2	Scienze Infermieristiche	1
B01	Anatomia	1	Fisioterapista	1
B03	Lingua Inglese	1	Scienze Infermieristiche	2
C09	Informatica	2	Tecn. Prevenzione	2
A06	Matematica Applicata	2	Fisioterapista	2
D02	Fisica I	1	Tecn. Prevenzione	1
...
MATERIA				

La tabella MATERIA registra i dati relativi alle materie di insegnamento.

un semplice esempio di database

La tabella DOCENTE registra i dati relativi ai docenti che insegnano nell'ateneo

Codice	Cognome	Nome	DataNascita	DataAssunzione	Residenza
DM 001	De Giorgis	Giovanni	13/07/1960	03/06/1998	Catanzaro
FT 002	Perli	Piero	02/12/1955	12/03/1987	Catanzaro
KR 003	Tripolino	Alfredo	07/05/1962	04/09/1990	Cosenza
...
DOCENTE					

CodDocente	CodMateria	MonteOre
DM 001	C 09	70
DM 001	D 02	55
FT 002	B 01	80
KR 003	A 01	75
...
INSEGNA		

La tabella INSEGNA merita un po' di attenzione, in quanto registra le interazioni tra i docenti dell'ateneo e le materie di insegnamento, dal momento che ci dice quale docente insegna cosa e per quante ore all'anno.

In questo, funziona come la tabella ESAME che, registrando i dati relativi ad un esame, registra di fatto le interazioni tra gli studenti e le materie di insegnamento, interazioni che si concretizzano nel momento in cui uno studente sostiene un esame in una determinata materia.

un semplice esempio di database

Da notare comunque come le tabelle non siano del tutto indipendenti tra loro; esistono infatti una serie di *associazioni logiche* tra i dati registrati che ci consentono di effettuare richieste di informazioni anche piuttosto complesse, associazioni rese possibili proprio dalle tabelle ESAME e INSEGNA, in particolare, dalla presenza in queste tabelle, dei codici identificativi di studenti, materie e docenti.

un semplice esempio di database

Volendo ad es. stabilire quali studenti abbiano superato l'esame di "Chimica Generale" con un voto superiore a 25/30, dovremmo innanzitutto controllare il codice della materia in questione verificandolo dalla tabella MATERIA (cioè A01),

Il codice dell'esame "Chimica Generale" è "A01".



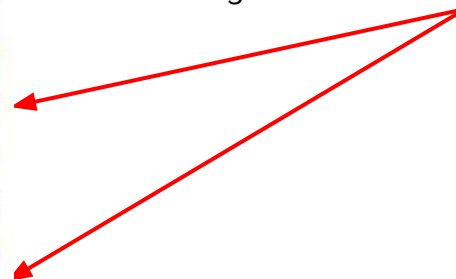
Codice	Denominazione	AnnoCorso	CorsoLaurea	Semestre
A 01	Chimica Generale	2	Scienze Infermieristiche	1
B 01	Anatomia	1	Fisioterapista	1
B 03	Lingua Inglese	1	Scienze Infermieristiche	2
C 09	Informatica	2	Tecn. Prevenzione	2
A 06	Matematica Applicata	2	Fisioterapista	2
D 02	Fisica I	1	Tecn. Prevenzione	1
...
MATERIA				

un semplice esempio di database

Quindi selezionare nella tabella ESAME tutte le righe che riportano come codice della materia il valore A01.

Matricola	CodMateria	DataEsame	Esito
8473	A01	09/04/2004	18
8841	A01	14/04/2004	27
8473	B01	01/06/2004	21
8473	C09	12/06/2004	30
7639	A01	23/06/2004	20
8841	D02	11/07/2004	25
...
ESAME			

Esami di “Chimica Generale” sostenuti dagli studenti.

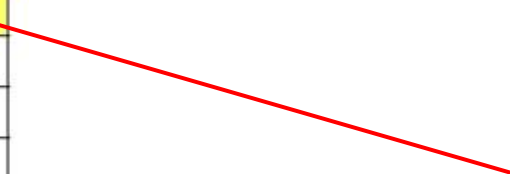


NB: Questa tabella ci da anche informazioni circa il voto conseguito e lo studente (attraverso la matricola) che ha sostenuto quel particolare esame.

un semplice esempio di database

Analizzando la tabella, possiamo ricavare le matricole dei soli studenti che hanno ottenuto una votazione superiore a 25/30 (basta verificare i valori registrati nella colonna “Esito”).

Matricola	CodMateria	DataEsame	Esito
8473	A01	09/04/2004	18
8841	A01	14/04/2004	27
8473	B01	01/06/2004	21
8473	C09	12/06/2004	30
7639	A01	23/06/2004	20
8841	D02	11/07/2004	25
...
ESAME			




L'unico studente che ha superato l'esame di “Chimica Generale” con un voto superiore a 25/30 è quello con matricola “8841”.

un semplice esempio di database

A questo punto, avendo a disposizione il numero di matricola, si possono ricavare anche gli altri dati anagrafici dello studente corrispondente prelevandoli dalla tabella STUDENTE.

Lo studente in questione è: Ugo Ragusa (matr.8841), nato il 04/05/1980, residente a Crotone, iscritto dal 04/10/2003, ecc.



Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1980	02/09/2003	Catanzaro	Sc. Infermierist.	1
7639	Criniti	Anna	02/09/1975	12/10/2002	Soverato	Fisioterapista	2
8841	Ragusa	Ugo	04/05/1980	04/10/2003	Crotone	Tecn. Prevenz.	1
...
STUDENTE							

un semplice esempio di database

Le operazioni sopra descritte possono essere eseguite solo in virtù del fatto che molti codici identificativi (come la matricola per lo studente e il codice dell'esame) sono replicati anche in altre tabelle preposte alla registrazione di dati di natura differente.

In pratica registrare nella tabella ESAME la matricola dello studente ed il codice dell'esame sostenuto, permette di stabilire una sorta di “collegamento logico” tra il singolo studente e ciascun esame da lui sostenuto, ogniqualevolta questi registra un esame nella tabella ESAME.

un semplice esempio di database

Abbiamo visto dunque come un DB possa essere strutturato in un **insieme integrato di archivi** (*tabelle*) collegati in qualche modo tra loro.

All'interno degli archivi i dati vengono organizzati in unità logiche riferite ad un unico soggetto e denominate **record** (*registrazione*); le singole informazioni che compongono il record si chiamano invece **campi**.

Record

Campi

Matricola	Cognome	Nome	DataNascita	DataIscrizione	Residenza	CorsoLaurea	AnnoCorso
8473	Baldieri	Marco	12/07/1980	02/09/2003	Catanzaro	Scienze Infermierist.	1
7639	Criniti	Anna	02/09/1975	12/10/2002	Soverato	Fisioterapista	2
8841	Ragusa	Ugo	04/05/1980	04/10/2003	Crotone	Tecn. Prevenzione	1
...
STUDENTE							

un semplice esempio di database

Ad es. l'archivio anagrafico dei soci di un club potrebbe essere organizzato come mostrato in figura, dove ogni socio è un singolo record, mentre ogni intestazione di colonna rappresenta un campo:

	COGNOME	TIT.	NOME	DATA NASCITA	CONIUGATO	TELEFONO
REC. 01	ROSSI	AVV.	MARIO	12/04/1956	VERO	47367823
REC. 02	CONTE	RAG.	MARCO	10/09/1962	FALSO	45009876
REC. 03	MANFREDI		ENZO	09/02/1960	VERO	12985558
REC. 04	PETRUZZI	DOTT.	VIOLA	23/06/1943	VERO	77823009
...						
REC. N	ALBERTI		CARLO	12/03/1971	FALSO	76904451

Si definisce ***tracciato del record*** l'insieme dei campi che compongono ciascun record di una stessa tabella; ad es:

SOCI CLUB: (Cognome, Titolo, Nome, Data nascita, Coniugato, Telefono)