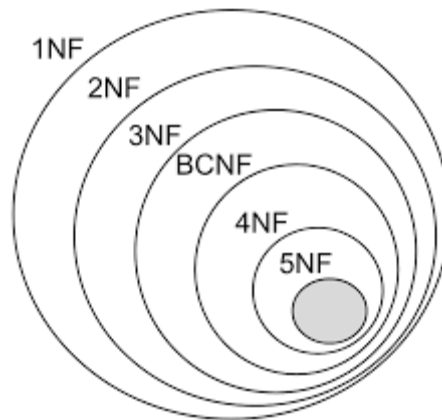


# normalizzazione

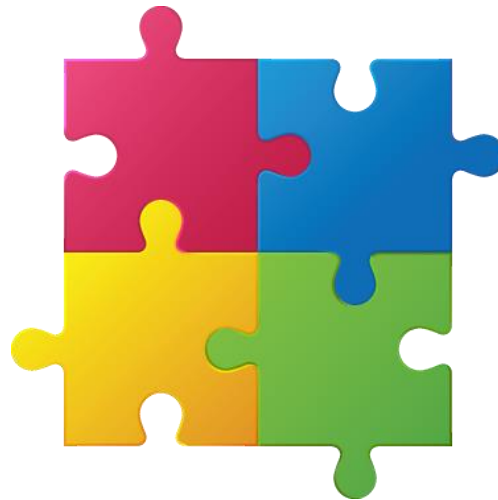
informatica

- la **normalizzazione** è un **procedimento** che permette di eliminare la ridondanza
- la **ridondanza** dei dati
  - comporta **spreco di memoria** causato dalla memorizzazione multipla della stessa informazione
  - può portare il database in una situazione di **incoerenza**
- si ha **incoerenza** quando **lo stesso campo ha valori diversi** in tabelle diverse
  - si può verificare quando le tabelle non sono aggiornate o quando l'aggiornamento non è stato effettuato correttamente

- esistono vari **livelli di normalizzazione (forme normali)** che certificano la qualità dello schema del database

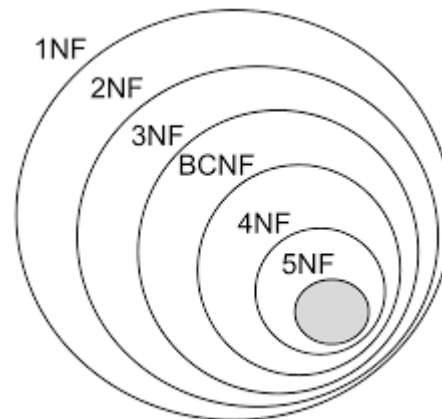


- se una relazione presenta **più concetti** tra loro indipendenti, la si **decompone** in relazioni più piccole, una per ogni concetto



- la ***dipendenza funzionale*** è un ***vincolo*** di integrità che descrive legami di tipo funzionale tra gli attributi di una relazione
- se un attributo ***a*** **determina** un altro attributo ***b*** allora esiste una **dipendenza funzionale tra *a* e *b***
- definizione:
  - data una relazione ***X*** e due sottoinsiemi di attributi ***Y*** e ***Z*** di ***X***, esiste una **dipendenza funzionale** tra ***Y*** e ***Z*** se **per ogni coppia** di tuple ***t1*** e ***t2*** aventi gli stessi valori sugli attributi ***Y***, risulta che ***t1*** e ***t2*** hanno gli **stessi valori** anche sugli attributi ***Z***

- una base dati è in **prima forma normale** se e solo se:
  - **non** presenta **attributi multipli**
  - esiste una **chiave primaria**
    - (esiste un insieme di attributi, che identifica in modo univoco ogni tupla della relazione)



*non in 1NF*

**Voti**

| <u>Matricola</u> | Studente | <u>Materia</u> | Voto                |
|------------------|----------|----------------|---------------------|
| 0000-000-01      | Pietro   | Basi di Dati   | 1 sem, B ; 2 sem, F |
| 0000-000-02      | Pietro   | Basi di Dati   | 1 sem, A ; 2 sem, A |
| 0000-000-03      | Sara     | Basi di Dati   | 1 sem, B ; 2 sem, A |

*in 1NF*

**Voti**

| <u>Matricola</u> | Studente | <u>Materia</u> | <u>Semestre</u> | Voto |
|------------------|----------|----------------|-----------------|------|
| 0000-000-01      | Pietro   | Basi di Dati   | 1               | B    |
| 0000-000-01      | Pietro   | Basi di Dati   | 2               | F    |
| 0000-000-02      | Pietro   | Basi di Dati   | 1               | A    |
| 0000-000-02      | Pietro   | Basi di Dati   | 2               | A    |
| 0000-000-03      | Sara     | Basi di Dati   | 1               | B    |
| 0000-000-03      | Sara     | Basi di Dati   | 2               | A    |

## Computer

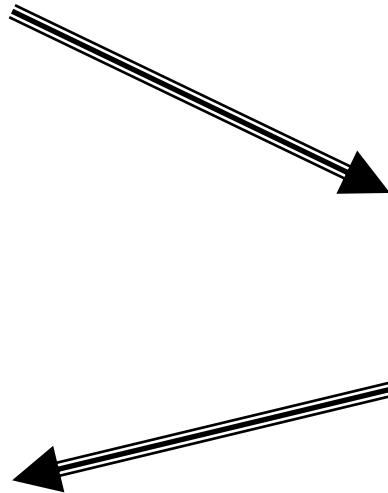
| Codice | Descrizione              | Processore            |
|--------|--------------------------|-----------------------|
| 1      | Desktop base             | Intel, CoreDuo, 2GHz  |
| 2      | Desktop medio            | AMD, Ahtlon XP, 3GHz  |
| 7      | Desktop fascia alta      | Intel, CoreDuo, 2GHz  |
| 8      | Portatile wireless medio | Transmeta, TX12, 1GHz |

## Computer

| Codice | Descrizione              | CodiceProcessore |
|--------|--------------------------|------------------|
| 1      | Desktop base             | P1               |
| 2      | Desktop medio            | P2               |
| 7      | Desktop fascia alta      | P1               |
| 8      | Portatile wireless medio | P3               |

## Processori

| CodiceProcessore | Marca     | Tipo      | Velocità |
|------------------|-----------|-----------|----------|
| P1               | Intel     | CoreDuo   | 2GHz     |
| P2               | AMD       | Ahtlon XP | 3GHz     |
| P1               | Intel     | CoreDuo   | 2GHz     |
| P3               | Transmeta | TX12      | 1GHz     |

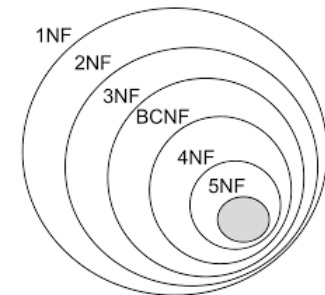


## Processori

| CodiceProcessore | Marca     | Tipo      | Velocità |
|------------------|-----------|-----------|----------|
| P1               | Intel     | CoreDuo   | 2GHz     |
| P2               | AMD       | Ahtlon XP | 3GHz     |
| P3               | Transmeta | TX12      | 1GHz     |



- una relazione è in **seconda forma normale** quando è in **1NF** e tutti i campi **non chiave** dipendono funzionalmente **dall'intera chiave** composta e **non da una parte di essa**
- la seconda forma normale **elimina la dipendenza parziale** degli attributi dalla chiave e riguarda il caso di relazioni con **chiavi composte**, cioè formate da più attributi



| Codice_atleta | Numero_gara | Nome_atleta | tempo |
|---------------|-------------|-------------|-------|
| 3455          | 1           | Maurizio    | 36    |
| 3455          | 2           | Maurizio    | 32    |
| 3320          | 1           | Luigi       | 38    |
| 3320          | 2           | Luigi       | 39    |
| 3320          | 3           | Luigi       | 33    |

la **chiave** è **composta**, in quanto il solo codice non basta per identificare il tempo che è realizzato in gare diverse.

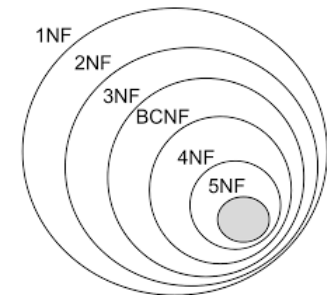
**Nome\_atleta** dipende solo dall'attributo Codice\_atleta (*solo da una parte della chiave*)

| Codice_atleta | Numero_gara | tempo |
|---------------|-------------|-------|
| 3455          | 1           | 36    |
| 3455          | 2           | 32    |
| 3320          | 1           | 38    |
| 3320          | 2           | 39    |
| 3320          | 3           | 33    |

| Codice atleta | Nome atleta |
|---------------|-------------|
| 3455          | Maurizio    |
| 3320          | Luigi       |

la relazione viene “spezzata” in due relazioni per eliminare l’anomalia

- una relazione si dice in *terza forma normale* quando è in **seconda** forma normale e **tutti** gli attributi **non-chiave** dipendono dalla chiave **soltanto**, ossia non esistono attributi che dipendono da altri attributi non-chiave



## Computer

| <b>Codice</b> | <b>Descrizione</b>       | <b>CodiceMemoria</b> | <b>Dimensione</b> | <b>Codice Processore</b> | <b>Velocità</b> |
|---------------|--------------------------|----------------------|-------------------|--------------------------|-----------------|
| 1             | Desktop base             | 20                   | 512MB             | P1                       | 2GHz            |
| 2             | Desktop medio            | 21                   | 1GB               | P2                       | 3GHz            |
| 7             | Desktop fascia alta      | 22                   | 2GB               | P3                       | 4GHz            |
| 8             | Portatile wireless medio | 25                   | 1GB               | P2                       | 3GHz            |

- Codice → Descrizione;
- CodiceMemoria → Dimensione;
- CodiceProcessore → Velocità.

## Computer

| <b>Codice</b> | <b>Descrizione</b>       | <b>CodiceMemoria</b> | <b>CodiceProcessore</b> |
|---------------|--------------------------|----------------------|-------------------------|
| 1             | Desktop base             | 20                   | P1                      |
| 2             | Desktop medio            | 21                   | P2                      |
| 7             | Desktop fascia alta      | 22                   | P3                      |
| 8             | Portatile wireless medio | 25                   | P2                      |

## Processori

| <b>CodiceProcessore</b> | <b>Velocità</b> |
|-------------------------|-----------------|
| P1                      | 2GHz            |
| P2                      | 3GHz            |
| P1                      | 2GHz            |
| P3                      | 1GHz            |

## Memorie

| <b>CodiceMemoria</b> | <b>Dimensione</b> |
|----------------------|-------------------|
| 20                   | 512MB             |
| 21                   | 1GB               |
| 22                   | 2GB               |
| 25                   | 1GB               |

| <u>CodImpiegato</u> | Nome    | Reparto | CapoReparto |
|---------------------|---------|---------|-------------|
| 1                   | Verdi   | Vendite | Rossi       |
| 2                   | Bianchi | Vendite | Rossi       |

- nell'esempio il campo CapoReparto *dipende funzionalmente* dal campo non-chiave Reparto
- la soluzione è quella di *scomporre* in due relazioni eliminando il campo che ha dipendenza funzionale dalla relazione di partenza

- **Prezzo** è deducibile dai due attributi non chiave **Costo\_unitario** e **Quantità**
- quindi **Prezzo** va eliminato

| Codice_prodotto | Costo_unitario | Quantità | Prezzo |
|-----------------|----------------|----------|--------|
| 1040            | 23             | 2        | 46     |
| 567             | 100            | 6        | 600    |
| 99854           | 55             | 1        | 55     |