

# Software Visuali per Analisi Avanzate

Corso sostitutivo di Tirocinio



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

1

## Comunicazioni di servizio

- Ultimo giorno per ricevere gli elaborati di SaS e Tableau
- Link Slide del corso
  - <https://github.com/valeriovvv/Corso-Software-Visuali-per-Analisi-Avanzate>
- Attestati di partecipazione SaS via mail a breve



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

2

## Comunicazioni di servizio



- Download Tableau licenza full educational
  - <https://www.tableau.com/academic/students>
- Download Tableau Trial
  - <https://www.tableau.com/it-it/products/trial>
- E-learning Tableau gratis per 90 giorni:
  - <https://www.tableau.com/learn/training/elearning>
- Badge da mettere sul proprio profilo LinkedIn:
  - <https://www.tableau.com/learn/learning-paths/consumer>



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

3

## Installazione di MySQL



- Installazione locale sul proprio PC Windows e Mac:
  - OK se il PC è abbastanza performante, si disponi di spazio e dei permessi
  - <https://www.apachefriends.org/it/download.html>
  - Installare seguendo le istruzioni (se viene richiesto, concedere accesso solo su reti private)
  - Avviare Xampp Control Panel
  - Avviare (start) Apache e Mysql
  - Aprire il browser sul link: <http://localhost/phpmyadmin/>
- Utilizzo in Cloud (nessuna installazione sul pc):
  - <https://www.db4free.net/>
  - Creare un account su: <https://www.db4free.net/signup.php>
  - Collegarsi al link: <https://www.db4free.net/phpMyAdmin/>



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

4

## Calendario del corso

Giorno	Durata	Argomenti
Martedì 5/5 16.00 – 18.00	2 ore	Presentazione del Corso Dati Strutturati, semi-strutturati e non strutturati Software da installare per il corso
Sabato 9/5 9.30 - 13.30	4 ore	Database relazionali MySQL Database, Tabelle, Righe, Colonne Linguaggio SQL: Introduzione, Select, Where, Join
Martedì 12/5 16.00 – 18.00	2 ore	Linguaggio SQL: Creazione, modifica, eliminazione di tabelle
Sabato 16/5 9.30 - 13.30	4 ore	Linguaggio SQL: Query di aggregazione Ottimizzazione delle Query: Gli indici Operazioni sui database Introduzione ai permessi (Grant)
Martedì 19/5	2 ore	Analitica Avanzata in ambienti Big Data Introduzione all'ecosistema Hadoop Apache Spark
Sabato 23/5	4 ore	Databricks e suoi componenti Utilizzo dell'SQL per interrogare Big Data Analitica avanzata con Databricks e R
Martedì 26/5	2 ore	Machine Learning con Databricks e R
Sabato 30/5	4 ore	Introduzione al Deep Learning, frontiera dell'analitica avanzata Conclusione del corso

Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020



5

## Tecnologie



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

6

## Sommario

Gartner Magic Quadrant for RDBMS

I database relazionali

Il linguaggio SQL

Esercitazione

Gartner  
Magic  
Quadrant

- **Magic Quadrant (MQ)** è una serie di ricerche di mercato pubblicate dalla società di consulenza **Gartner** basati su metodi di analisi proprietari dei dati per mostrare le tendenze del mercato.
- Le analisi sono condotte per diversi settori tecnologici specifici e vengono periodicamente aggiornate.
- Gartner valuta i fornitori in base a due criteri: **completeness of vision** (completezza della visione) e **ability to execute** (capacità di esecuzione)
- Il report di Gartner si divide in 4 quadranti

## Gartner Magic Quadrant

- **Leaders** – In questo quadrante sono presenti i vendor che hanno punteggio più alto per Completeness of Vision e Ability to Execute. Hanno quote di mercato, credibilità e le capacità di marketing e di vendita necessarie a guidare la tecnologia al successo. Questi vendor dimostrano chiara comprensione delle necessità del mercato, hanno un pensiero innovativo e piani ben definiti.
- **Challengers** – I vendor di questo quadrante sono presenti nel mercato ed hanno una Ability to Execute buona, tanto da costituire una seria minaccia per i venditori nel quadrante Leader. Hanno prodotti validi, posizione di mercato e risorse sufficientemente credibili per sostenere la crescita continua. Hanno buona redditività finanziaria, ma non hanno le dimensioni e l'influenza dei venditori nel quadrante Leader.

## Gartner Magic Quadrant

- **Visionaries** – Un fornitore nel quadrante Visionaries offre prodotti innovativi che affrontano bene i problemi degli utenti finali a livello operativo o finanziario, ma non ha ancora dimostrato la capacità di acquisire quote di mercato o redditività sostenibile. I venditori visionari sono spesso aziende private e obiettivi di acquisizione per aziende più grandi e affermate. La probabilità di essere acquisiti spesso riduce i rischi associati all'adozione dei loro sistemi.
- **Niche Players** – I vendors di questo quadrante sono spesso focalizzati su segmenti di mercato o esigenze verticali specifiche. Questo quadrante può anche includere venditori che stanno riadattando loro prodotti esistenti per entrare nel mercato, o venditori più grandi che hanno difficoltà a far affermare la propria visione.

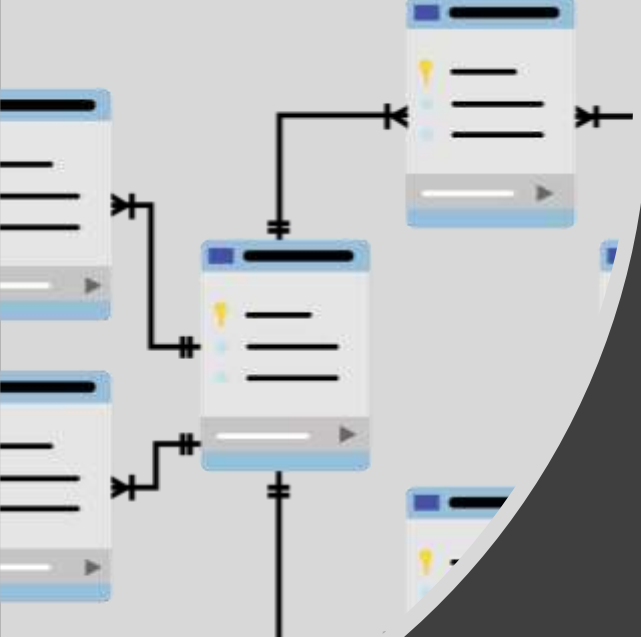
## Magic Quadrant for Operational Database Management Systems

Data and analytics leaders evaluating OPDBMS options must balance their current and future needs in an increasingly competitive, “cloud-first” market. Nontraditional vendors offer new opportunities for global enterprises, which can use this Magic Quadrant to make better modernization choices.




<https://www.gartner.com/doc/reprints?id=1-1YAE9AY1&ct=200206&st=sb&signin=640614711143974a616dfab8ab01ef7d>

11



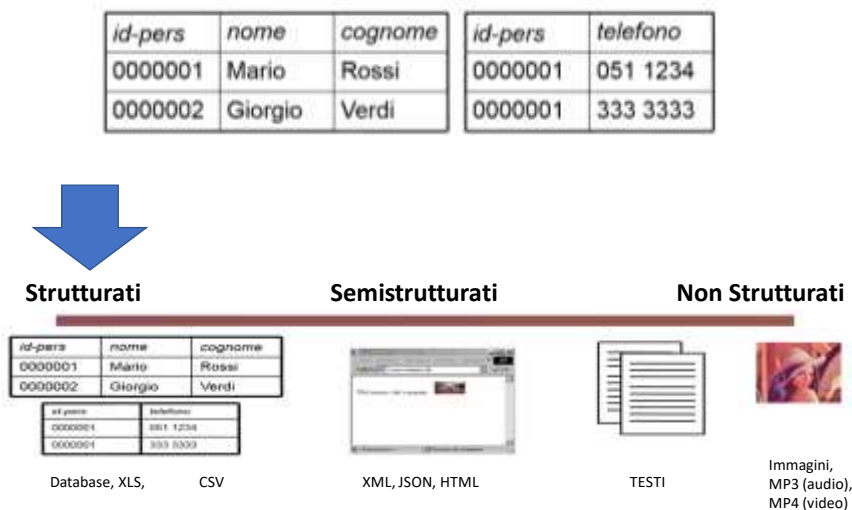
# I Database Relazionali



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

12

## Dati Strutturati



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

13

## Dati Strutturati

- Dati in posizione fissa all'interno di record con un tipo ben definito (es. testo, numero, data, ecc.)
- La struttura dei dati (definizione o metadati) è ben definita e rigida
- Metadati e Dati sono separati
  - Es. Nomi e tipo delle colonne di un file XLS e dati contenuti nelle celle
  - In un Database, tabelle, colonne e tipi di dato e dati contenuti
  - Nomi delle colonne di un CSV e valori delle righe
- Non è possibile avere strutture gerarchiche
  - Una cella di XLS può contenere un testo, un numero, non un elenco
- Alcuni esempi: Database Relazionali, Fogli di calcolo, CSV

14

## Dati Strutturati

- Colonne in posizione fissa
- Netta separazione tra dati e metadati
- Tipi di dato ben definiti
- Non è possibile avere strutture gerarchiche

id-pers	nome	cognome	id-pers	telefono
0000001	Mario	Rossi	0000001	051 1234
0000002	Giorgio	Verdi	0000001	333 3333

	A	B	C	D	E	F
1	Nome	Cognome	Data di Nascita	Media		
2	Mario	Rossi	13/01/1998	28,7		
3	Maria	Verde	24/10/2001	27,3		
4	Luisa	Bianchi	10/09/2000	29		
5	Oreste	Arancio	11/07/1999	24,6		
6						
7						

data;stato;codice\_regione;denominazione\_regione;lat;long;ricoverati\_con\_sintomi  
 2020-02-24T18:00:00;ITA;17;Basilicata;40.63947052;15.80514834;0  
 2020-02-24T18:00:00;ITA;18;Calabria;38.90597598;16.59440194;0  
 2020-02-24T18:00:00;ITA;15;Campania;40.83956555;14.25084984;0  
 2020-02-24T18:00:00;ITA;08;Emilia-Romagna;44.49436681;11.3417208;10

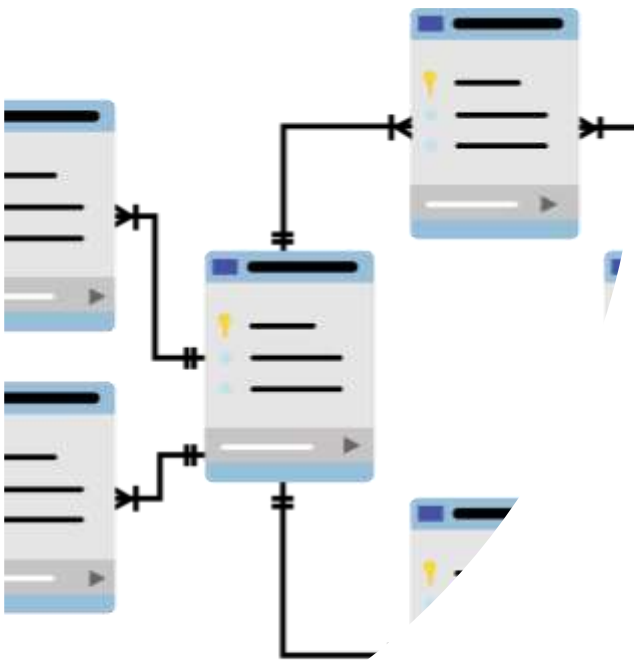
Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020



## Database Relazionali

- Introduzione ai Database Relazionali
  - <https://www.slideshare.net/ninbarslideshare/database-introduzione>
- Approfondimento
  - <http://www.diag.uniroma1.it/~catarci/basidatGEST.html>





# Introduzione a MySQL

MySQL

- MySQL è un database relazionale
- Funziona come processo di sistema
- E' possibile dare comandi a MySQL attraverso una console a riga di comando
- Oppure tramite PhpMyadmin, una console grafica scritta in linguaggio PHP

## Accedere a PhpMyAdmin



19

- Accesso Xampp:
  - Nel browser accedere all'indirizzo:  
<http://localhost/phpmyadmin>
- Accesso db4free
  - <https://www.db4free.net/phpMyAdmin>
  - Inserire userid e password ottenuti in fase di registrazione
- Esplorazione dell'interfaccia grafica

Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

## Accedere a PhpMyAdmin



20

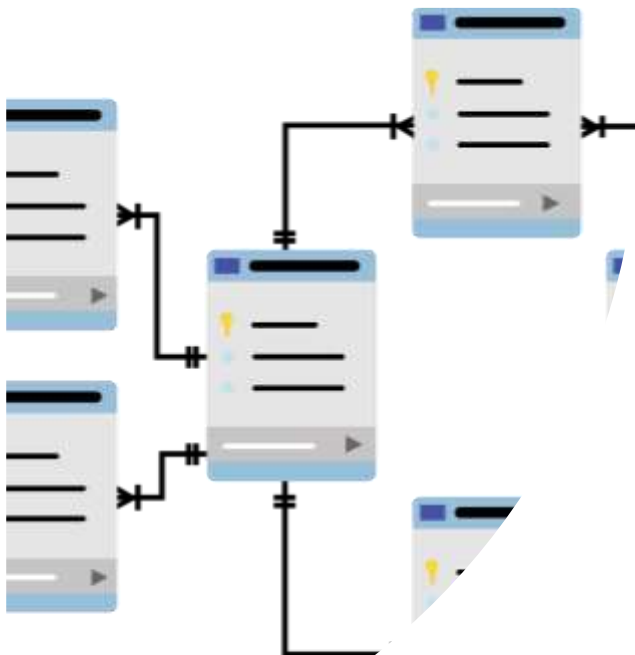
- Creazione di un nuovo database (solo per chi ha XAMPP)
- Creazione tabella da interfaccia grafica
- Inserimento dati nelle tabelle
- Visualizzare i dati inseriti
- La prima query
- Designer
- Esportare i dati
- Importare i dati

Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

## Tabella da creare

maternita		persone		
madre	figlio	nome	eta	reddito
Luisa	Maria	Andrea	27	21
Luisa	Luigi	Aldo	25	15
Anna	Olga	Maria	55	42
Anna	Filippo	Anna	50	35
Maria	Andrea	Filippo	26	30
Maria	Aldo	Luigi	50	40
		Franco	60	20
		Olga	30	41
		Sergio	85	35
		Luisa	75	87

- Definizione della chiave primaria
- Aggiungere un campo:
  - ID
  - Di tipo intero
  - Con flag A\_I (autoincrement)



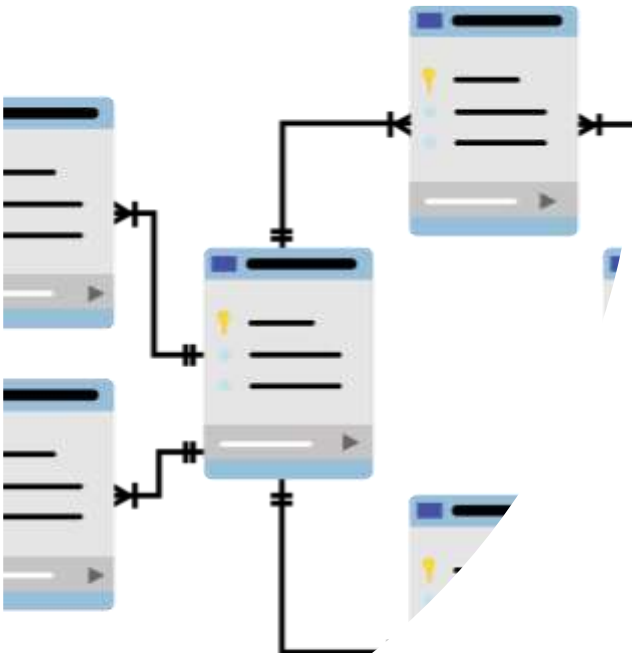
## Il Linguaggio SQL



## SQL

- Il linguaggio SQL
  - <http://www.dis.uniroma1.it/~disanzo/data/Basi%20di%20dati%202008-2009/3%20-%20Sql.pdf>
  - **Partiamo dalla slide 12**
- Approfondimenti
  - [http://ai-nlp.info.uniroma2.it/basili/didattica/DB\\_18\\_19/003\\_CAP3\\_IntroModRelazionale\\_2014\\_15\\_parte1.pdf](http://ai-nlp.info.uniroma2.it/basili/didattica/DB_18_19/003_CAP3_IntroModRelazionale_2014_15_parte1.pdf)
  - [http://ai-nlp.info.uniroma2.it/basili/didattica/DB\\_18\\_19/CAP5-LezioneSQL\\_aa2018\\_19.pdf](http://ai-nlp.info.uniroma2.it/basili/didattica/DB_18_19/CAP5-LezioneSQL_aa2018_19.pdf)
  - [http://www.diag.uniroma1.it/~catarci/slides/BD/4-SQL\\_Structured\\_Query\\_Language.pdf](http://www.diag.uniroma1.it/~catarci/slides/BD/4-SQL_Structured_Query_Language.pdf)
  - [http://www.diag.uniroma1.it/~catarci/slides/BD/6-SQL\\_Interrogazioni\\_Nidificate.pdf](http://www.diag.uniroma1.it/~catarci/slides/BD/6-SQL_Interrogazioni_Nidificate.pdf)

23



## Esercitazione

24

## Esercitazione 1

1. Creare una tabella con i seguenti campi:
  - ID (autoincrement)
  - Codice Corso (Varchar 10)
  - Titolo Corso (Varchar 50)
  - Anno accademico (Varchar 10)
  - CFU (intero)
2. Inserire 5 corsi da interfaccia grafica (con CFU: 3, 6, 9, 12)
3. Creare una query che recuperi tutti i campi
4. Creare una query che recuperi Codice Corso e CFU (rinominato in Crediti\_Formativi)
5. Creare una query che recuperi Codice Corso e CFU ordinati per CFU decrescente
6. Creare una query che ritorni il codice ed il titolo dei corsi che hanno un numero di CFU <6
7. Creare una query che ritorni il codice ed il titolo dei corsi che hanno un numero di CFU compresp fra 6 e 9.
8. Creare un elenco di tutti gli studenti che hanno il nome che inizia con la lettera «A»

## Esercitazione 2

- Recuperare il Json creato nella lezione precedente...
- Creare l'equivalente del modello Json su Mysql
- Come hai risolto la gerarchia presente nel Json utilizzando il modello relazionale?
- Quante tabelle hai creato? Hai aggiunto qualche campo?
- Creare una query che recupera tutti gli studenti
- Creare una query che recupera Matricola, Esame, Data, Voto
- Creare una query che ritorna i nomi univoci degli studenti
- Creare una query che calcola: voto minimo, massimo e medio (in generale per tutti gli studenti)

## Esercitazione 2

- Recuperare il Json creato nella lezione precedente...



```
{
  "Studenti": [
    {
      "Nome": "Antonio",
      "cognome": "Rossi",
      "matricola": "0013424324",
      "esami": [
        {
          "Esame": "Statistica",
          "Data": "10/01/2020",
          "Voto": 30
        }
      ]
    }
  ],
  {
    "Nome": "Marco",
    "cognome": "Verdi",
    "matricola": "0013425655",
    "esami": [
      {
        "Esame": "Statistica",
        "Data": "10/01/2020",
        "Voto": 29
      }, {
        "Esame": "Economia",
        "Data": "11/11/2020",
        "Voto": 28
      }
    ]
  }
]
```



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

27

## Esercizio 3

- Aggiungere alla tabella Esami il campo codice corso ed eliminare il campo Corso
- Perché pensi sia utile eliminare il campo corso?

studenti studente	
Nome	varchar(20)
Cognome	varchar(20)
Matricola	varchar(10)

studenti corsi	
ID	int(11)
codice_corso	varchar(10)
titolo_corso	varchar(50)
anno_accademico	varchar(10)
cfu	int(11)

studenti esami	
ID	int(11)
codice_corso	varchar(5)
data	date
voto	int(11)



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

28

## Esercizio 3

- Assicurati di avere almeno 5 studenti, 10 esami, 5 corsi
- Crea le seguenti query:
  - Elenco anni accademici (ogni anno accademico deve comparire una sola volta)
  - Elenco degli studenti in ordine alfabetico
  - Elenco esami sostenuti da uno studente (a tua scelta) ordinati per Data esame con i seguenti campi: Nome, Cognome, Matricola, Titolo Corso, Anno accademico, Voto, data esame
  - Calcola la media dei voti per un singolo studente a tua scelta
  - Nome, Cognome, Matricola, Titolo Corso dello studente che ha sostenuto l'esame più recente
  - Esporta le tabelle create



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

29

## Esercizio 4

- Esporta le tabelle create nell'esercizio 3
- Collega Tableau al database dell'esercizio 3 e crea un dashboard che illustri le caratteristiche principali delle informazioni

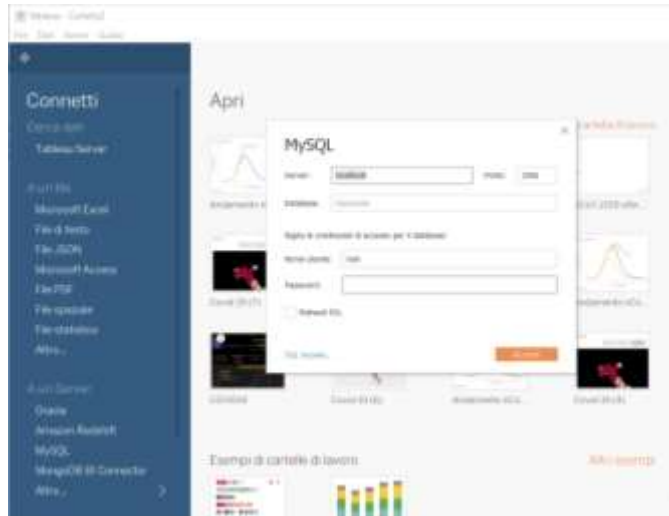


Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

30

## Collegare Tableau a Mysql (Xampp)

- Scaricare Tableau Educational o Trial
- Andare nelle connessioni
- Selezionare Mysql
- Server: localhost
- Database:
- Nome Utente: root



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020

## Collegare Tableau a Mysql (db4free)

- Scaricare Tableau Educational o Trial
- Andare nelle connessioni
- Selezionare Mysql
- Server: db4free.net
- Database:
- Nome Utente: <tuvo username>
- Password: <tua password>



Dott. Valerio Morfino – Sistemi Visuali per Analisi Avanzate a.a. 2019-2020



# Grazie per l'attenzione



<https://it.linkedin.com/in/valerio-morfino>



[vmorfino@unisannio.it](mailto:vmorfino@unisannio.it)

---