

Software Visuali per Analisi Avanzate

Corso sostitutivo di Tirocinio

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
- Rilevamento presenze: Quando richiesto, inserire in chat: Nome, Cognome e Matricola

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>

Installazione di MySQL



- Installazione locale sul proprio PC Windows e Mac:
 - OK se il PC è abbastanza performante, si dispon 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/>

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



Tecnologie



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.



As of July 2019

© Gartner, Inc



Dati Strutturati

<i>id-pers</i>	<i>nome</i>	<i>cognome</i>
0000001	Mario	Rossi
0000002	Giorgio	Verdi

<i>id-pers</i>	<i>telefono</i>
0000001	051 1234
0000001	333 3333



Strutturati

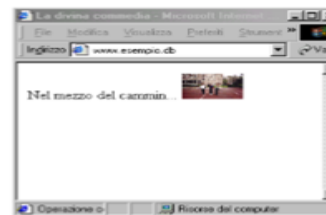
<i>id-pers</i>	<i>nome</i>	<i>cognome</i>
0000001	Mario	Rossi
0000002	Giorgio	Verdi

<i>id-pers</i>	<i>telefono</i>
0000001	051 1234
0000001	333 3333

Database, XLS,

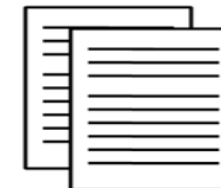
CSV

Semistruutturati

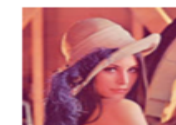


XML, JSON, HTML

Non Strutturati



TESTI



Immagini,
MP3 (audio),
MP4 (video)

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

Dati Strutturati

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

<i>id-pers</i>	<i>nome</i>	<i>cognome</i>
0000001	Mario	Rossi
0000002	Giorgio	Verdi

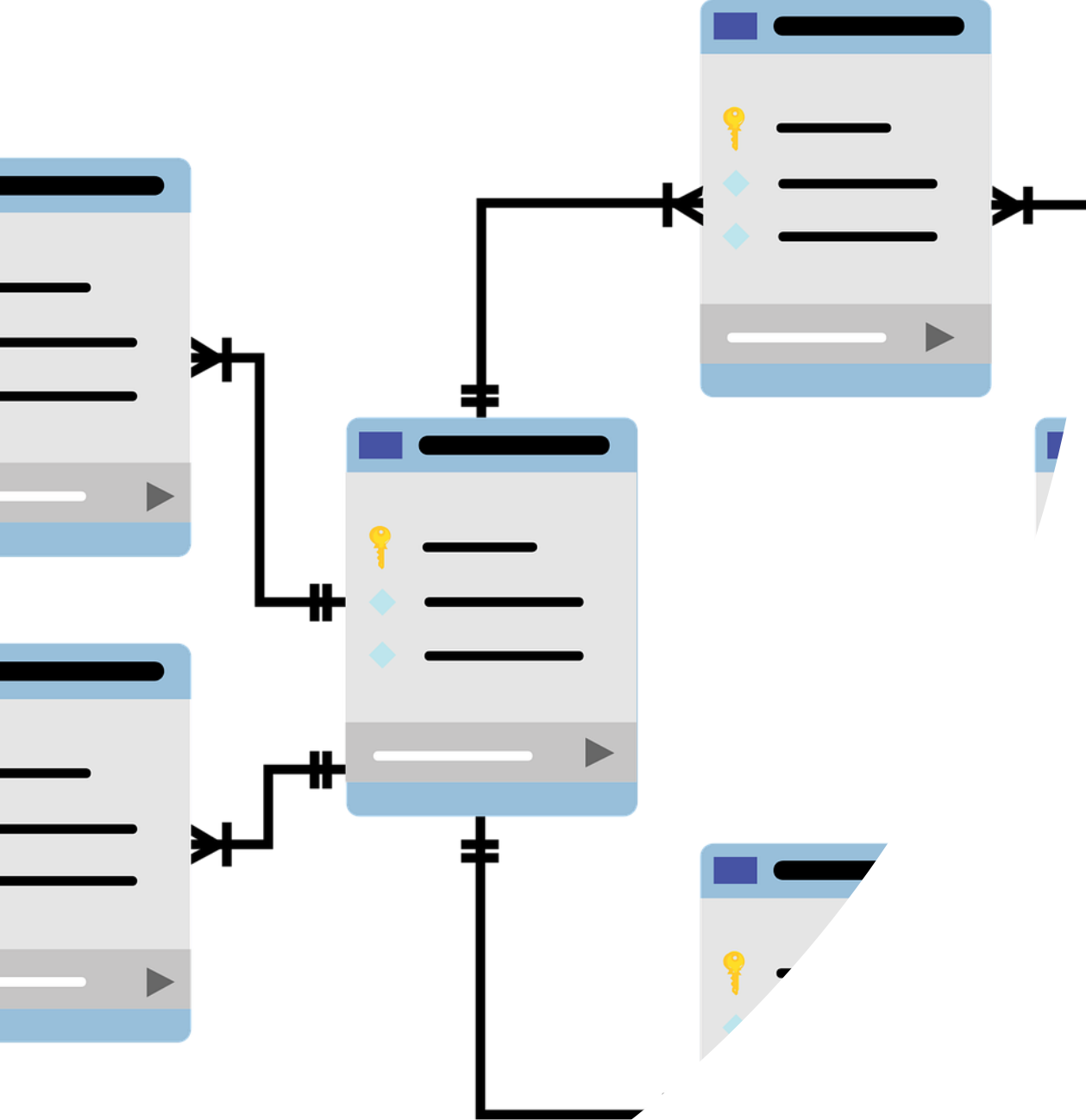
<i>id-pers</i>	<i>telefono</i>
0000001	051 1234
0000001	333 3333

	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

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 consolle a riga di comando
- Oppure tramite PhpMyadmin, una consolle grafica scritta in linguaggio PHP

Accedere a PhpMyAdmin

- 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

Accedere a PhpMyAdmin

- 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

Tabella da creare

maternita

madre	figlio
Luisa	Maria
Luisa	Luigi
Anna	Olga
Anna	Filippo
Maria	Andrea
Maria	Aldo

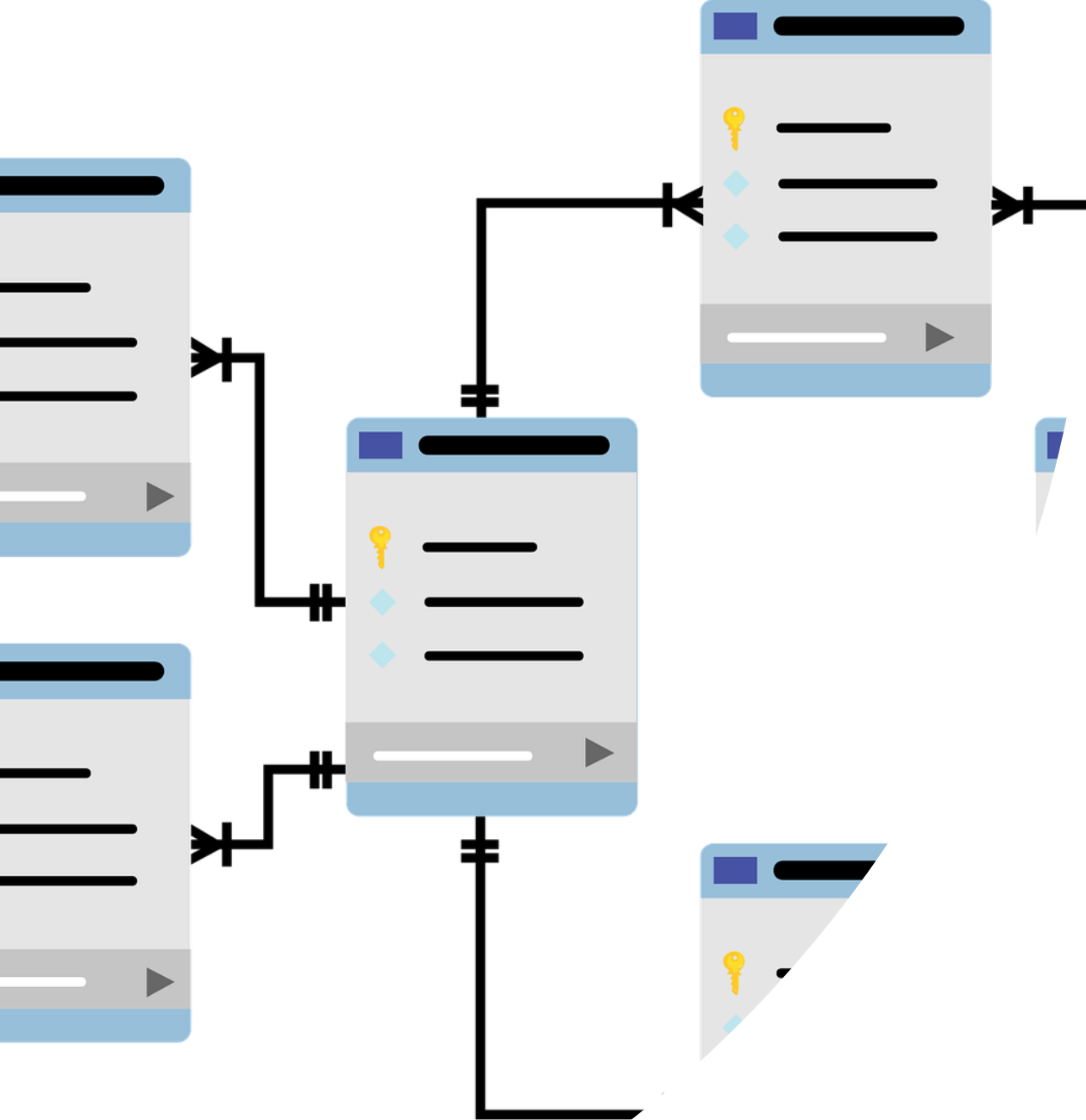
paternita

padre	figlio
Sergio	Franco
Luigi	Olga
Luigi	Filippo
Franco	Andrea
Franco	Aldo

persone

nome	eta	reddito
Andrea	27	21
Aldo	25	15
Maria	55	42
Anna	50	35
Filippo	26	30
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/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/6-SQL_Interrogazioni_Nidificate.pdf

Grazie per l'attenzione



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



vmorfino@unisannio.it
