

# Software Visuali per Analisi Avanzate

Corso sostitutivo di Tirocinio



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

1

## 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,
Martedì 12/5 16.00 – 18.00	2 ore	Linguaggio SQL: Inner Join, Left Join, Right Join, Order By
Sabato 16/5 9.30 - 13.30	4 ore	Linguaggio SQL: Query di aggregazione Cenni su Indici Cenni sulle DDL Esercitazione
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



2

## Comunicazioni di servizio

- Link Slide del corso
  - <https://github.com/valeriovvv/Corso-Software-Visuali-per-Analisi-Avanzate>
- Rilevamento presenze: Quando richiesto, inserire in chat: Nome, Cognome e Matricola
- Live Webinar SAS il 18/05 – Focus su Certificazione Machine Learning
- Modifica agenda e calendario del corso

### Proposta di Modifica

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,
Martedì 12/5 16.00 – 18.00	2 ore	Linguaggio SQL: Inner Join, Left Join, Right Join, Order By
Sabato 16/5 9.30 - 13.30	4 ore	Linguaggio SQL: Query di aggregazione Cenni su Indici Cenni sulle DDL Esercitazione finale SQL
Martedì 19/5	2 ore	Programmazione con Sas: Introduzione a SAS Studio Struttura di un programma SAS
Venerdì 22/5	4 ore	Programmazione SAS – Ospite docente di SAS Institute
Sabato 23/ o Martedì 26/5	4 ore	Programmazione con Sas SQL con SAS Esercitazione finale SAS
Martedì 26/5 oppure Venerdì 29/5	2 ore	Utilizzo di R e SQL in ambiente Big data: Spark e Databricks

## Tecnologie



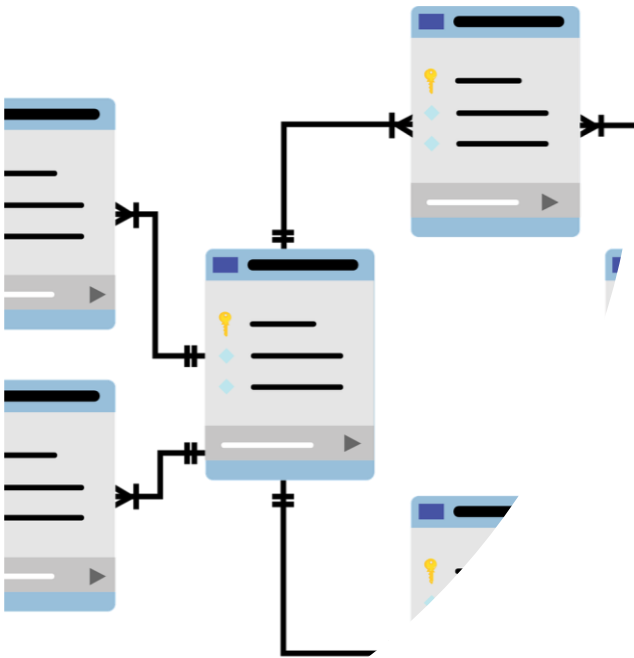
## Sommario

Linguaggio SQL

Query di Aggregazione

Cenni Indici e DDL

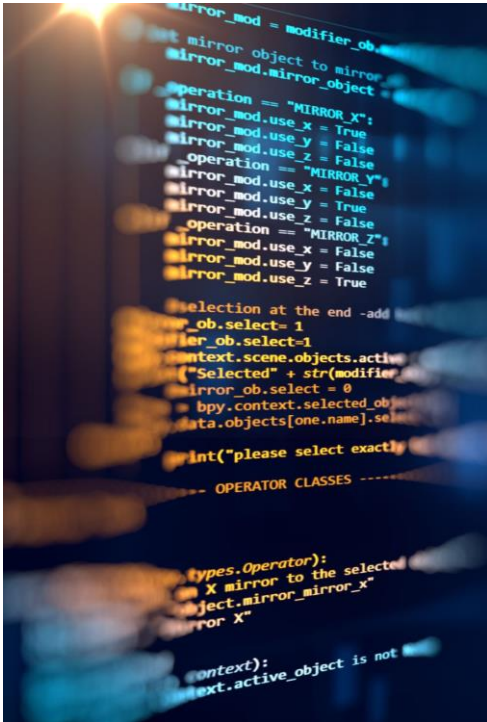
Esercitazione Finale



## Il Linguaggio SQL

### Importazione dati

- Scaricare il file:  
<https://github.com/valeriovvv/Corso-Software-Visuali-per-Analisi-Avanzate/blob/master/Modulo-2/impiegati.sql>
- Nella cartella c'è anche corso.sql con tutte le tabelle
- Due modi per importare lo script:
  1. Usando la funzione importa e selezionando il file
  2. Usando la funzione SQL:
    - Aprire il file con un text editor e copiare il contenuto
    - Copiare in SQL e cliccare esegui
- Osservare l'output



9

## SQL

- Il linguaggio SQL
  - <http://www.dis.uniroma1.it/~disanzo/data/Basi%20di%20dati%202008-2009/3%20-%20Sql.pdf>
  - **Partiamo dalla slide 32**
- 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)

Collegamento  
Tableau e  
Mysql



10

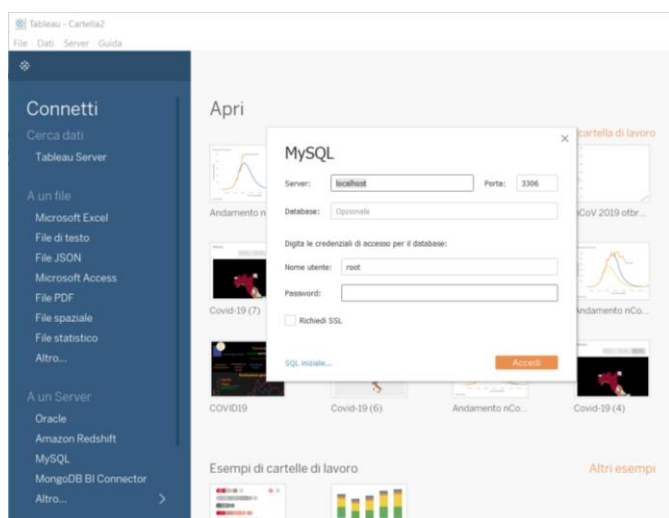
# Tableau Creator

- Tableau Public non permette di accedere a MySQL
- E' necessario utilizzare la versione commerciale di Tableau
- Download Tableau licenza full educational
  - <https://www.tableau.com/academic/students>
- Download Tableau Trial
  - <https://www.tableau.com/it-it/products/trial>

11

## Collegare Tableau a Mysql (Xampp)

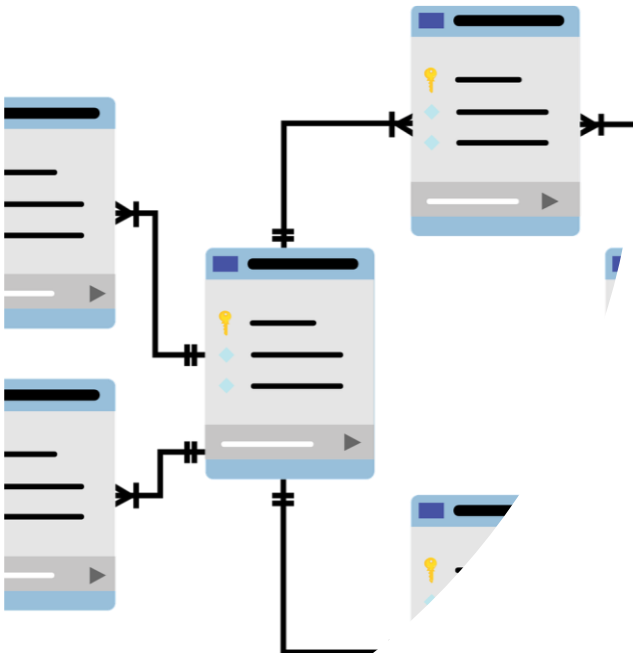
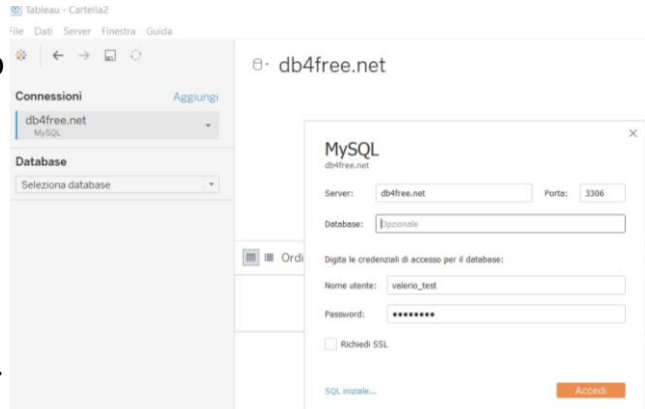
- Scaricare Tableau Educational o Trial
- Andare nelle connessioni
- Selezionare Mysql
- Server: localhost
- Database:
- Nome Utente: root



12

## 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>



## Esercitazione

## Esercizio Finale

- Database relativo alla posizione accademica degli studenti
- Lo schema è composto da 3 tabelle:
  - Studenti
  - Corsi
  - Esami

studenti studenti
nome : varchar(20)
cognome : varchar(20)
matricola : varchar(10)

studenti esami
ID : int(11)
codice_corso : varchar(10)
matricola : varchar(20)
data : date
voto : int(11)
lode : tinyint(1)

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

## Esercizio Finale – Definizione dello schema

1. Creare le tre tabelle su Mysql in base allo schema, complete di chiave primaria e tipi di dato corretti
2. Importare lo script studenti.sql presente su link GITHUB del corso
3. Modificare la tabella Studenti:
  - aggiungendo il campo «sesso» Varchar (1)
  - Valori ammessi: M oppure F
4. Modifica i dati della tabella inserendo il valore giusto per sesso
5. Esportare il database in modo da creare uno script SQL che contenga sia le DDL che le istruzioni per popolare il database



## Esercizio Finale – Analisi Studenti

1. Elencare tutti gli studenti (uomini e donne) in ordine alfabetico per cognome
2. Elencare le Studentesse ordinate per Matricola
3. Contare il numero totale degli studenti
4. Contare gli studenti raggruppati per Sesso: Quanti maschi e quante femmine
5. Creare una query che ritorni i nomi degli studenti senza duplicati
6. Contare il numero di nomi distinti degli studenti
7. Elencare i nomi degli studenti che si ripetono almeno due volte

## Esercizio Finale – Esami

1. Elencare codice, titolo e voto degli esami sostenuti dallo studente con Matricola A003
  1. `SELECT * FROM `studenti` inner join esami on  
studenti.matricola=esami.matricola where esami.matricola='A003'`
2. Stampare la media complessiva degli studenti dell'ateneo (il numero medio per tutti), il voto minimo e il voto massimo
3. Stampare la media, minimo e massimo dello studente con matricola A003
4. Stampare la media di tutti gli studenti (uno per riga)

## Esercizio Finale – Analisi sugli Esami

1. Elencare codice, titolo e voto degli esami sostenuti dallo studente con Matricola «A003»
2. Stampare la media dei voti degli studenti dell'ateneo, il voto minimo e il voto massimo (una sola riga con i dati di tutti gli studenti)
3. Stampare la media, minimo e massimo dello studente con matricola A003
4. Stampare la media, minimo e massimo per ciascuno studente (uno per riga)

## Esercizio Finale – Elenchi

1. Stampare l'elenco degli studenti che hanno sostenuto almeno un esame (nome, cognome, matricola)
2. Stampare l'elenco di tutti gli studenti e degli esami che hanno sostenuto (se ne hanno)
3. Stampare elenco degli studenti che non hanno sostenuto esami
4. Stampare elenco degli esami che non sono stati mai sostenuti da alcuno studente
5. Stampare Cognome, Matricola, Anno accademico, Media Esami per tutti gli studenti, ordinati per Anno accademico crescente e Media esami decrescente

## Esercizio Finale – Viz (Opzionale)

1. Connettere il database a Tableau
2. Creare una visualization per creare media esami per studente
3. Suddividere gli studenti in cluster
4. Analizzare gli esami considerando le medie dei voti ed il numero di volte che è stato sostenuto
5. Creare almeno un altro diagramma con i dati a disposizione che metta in risalto dati che sembrano interessanti
6. Creare una dashboard navigabile che raggruppi i dati

## Esercizio Finale – Json (Opzionale)

1. Studiare la rappresentazione JSON più adatta per «Studente»
2. Provare a creare il Json per uno studente qualsiasi

# Grazie per l'attenzione



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



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

---