

Nome docente	Francesco D. d'Ovidio
Corso di laurea	C.d.L in "Scienze Statistiche"
Insegnamento	BASI DI DATI
Anno accademico	2019-2020
Periodo di svolgimento	I semestre
Crediti formativi universitari (CFU)	6
Settore scientifico disciplinare	ING-INF/05
Pagine web docenti	https://www.uniba.it/docenti/dovidio-francesco

Pre-requisiti

Non si richiedono specifiche capacità e/o pre-requisiti di natura informatica

Per la corretta comprensione dell'ultima parte del corso, si richiede solamente conoscenza della statistica descrittiva e dei concetti base (media, varianza, indici statistici).

Conoscenze e abilità da acquisire (Obiettivi)

Il corso si propone innanzitutto di fornire i concetti fondamentali sulle basi di dati, sui linguaggi di interrogazione e gestione e sulle tecniche e metodi di progettazione. L'obiettivo è fornire agli studenti gli strumenti per la gestione e la progettazione di un sistema informativo che includa una base di dati.

Nella seconda parte del corso, gli studenti applicheranno le conoscenze acquisite anche all'interrogazione ed esplorazione statistica delle grandi basi di dati disponibili su server locali oppure su web-server. L'obiettivo finale è di fornire agli studenti i corretti strumenti statistici per un primo approccio ai Database di grandi dimensioni e ai Database non strutturati (i cosiddetti "Big Data").

Programma dettagliato

A. Fondamenti informatici

1. Introduzione alle Basi di Dati

Sistemi Informativi, Informazioni e dati; Basi di dati e sistemi di Gestione; Modelli di Dati; Linguaggi ed utenti delle BdD; Vantaggi e svantaggi dei DBMS.

2. Basi di Dati Relazionali: Modelli e linguaggi

Il modello relazionale: Strutture, Vincoli di Integrità;
Algebra e calcolo Relazionale: Algebra relazionale, Calcolo relazionale, Datalog.

3. Structured Query Language - SQL

Definizione dei dati in SQL; Interrogazioni in SQL; Manipolazione dei dati in SQL; Altre definizioni dei dati; Funzioni scalari; Controllo dell'accesso.

4. Progettazione di Basi di Dati

Metodologie e modelli per il progetto; Cenni di progettazione concettuale; La progettazione logica; La normalizzazione.

B. Fondamenti statistici

1. Gestione dei database con finalità di analisi statistica

Tecniche di integrazione di dataset rivenienti da fonti differenti: Record Linkage e Statistical Matching;
Tecniche di esplorazione statistica di grandi database; Introduzione al Data Mining.

2. Analisi statistiche di grandi database

Analisi statistiche non inferenziali; Tecniche statistiche applicabili all'analisi di grandi masse di dati;
Approccio Data Driven.

3. Dai Database ai Big Data: la Data Science

Dai database strutturati ai dati non strutturati; Il concetto di *NoSQL Relational Database Management System*; La nascita del Cloud; Campi di applicazione dell'analisi dei *Big Data*; La *Data Science*.

Riferimenti Bibliografici e Materiali didattici

- ✓ Atzeni P., Ceri S., Fraternali P., Paraboschi S., Torlone R. (2013). *Basi di Dati - Modelli e linguaggi di interrogazione*, 4/ed, McGraw-Hill.
- ✓ Curtin D. P., Foley K., Sen K., Morin C. (2016). *Informatica di base*, (6/ed), McGraw-Hill.
- ✓ Atzeni P., Ceri S., Fraternali P., Paraboschi S., Torlone R. (2014). *Basi di Dati*, McGrawHill.
- ✓ Marinuzzi F., Liciani M. (2016). *Basi di dati e big data: come estrarre valore dai propri dati*. Youcanprint.
- ✓ d'Ovidio F.D., *Lezioni* (versione PDF delle slides man mano presentate durante il corso, in formato compresso crittografato: la relativa pw è fornita agli studenti frequentanti al termine di ciascun argomento).

Organizzazione della didattica:

- Cicli interni di lezione: No
- Corsi integrativi: No
- Esercitazioni: **Sì**
- Seminari: No
- Attività di laboratorio: No
- Project work: **No**
- Visite di studio: No
- Disponibilità on-line di dispense e slide delle lezioni ad uso dei frequentanti: **Sì**

Modalità di erogazione delle attività formative:

Lezioni frontali riguardanti gli argomenti teorici ed esempi pratici, durante le quali vengono esposti dei problemi atti a sviluppare la capacità dello studente ad applicare la teoria nel contesto dei fenomeni reali. Saranno utilizzate metodologie di e-learning per coadiuvare lo studente nel percorso di studi, utilizzando un portale web “ad hoc” per migliorare l’interazione docente-studente e studente-studente, nonché ottimizzare il processo di reperimento dei materiali e di supporto alla formazione

Modalità di accertamento delle conoscenze:

- Non sono previsti esoneri e prove intermedie durante il corso.
- L’esame di profitto prevede una discussione teorico-pratica degli argomenti trattati nel corso per la verifica (possibilmente, la parte pratica sarà svolta separatamente rispetto a quella teorica).