



Corso di Informatica, Università degli studi di Udine

Laboratorio di Basi di Dati (2023-2024) – Lezione 1 –

Luca Geatti

luca.geatti@uniud.it

20 Ottobre 2023



- **Nome:** Luca Geatti
 - RTD-a @ Dipartimento di Scienze Matematiche, Informatiche e Fisiche
- **E-Mail:** luca.geatti@uniud.it
- **Website:** <https://users.dimi.uniud.it/~luca.geatti/>
- **Ufficio:** ... mandatemi una mail che ve lo spiego ...
- **Ricevimento:**
 - il Venerdì 14:00-16:00
 - sono molto flessibile sull'orario
 - meglio su appuntamento
- **Annunci:** su Teams
- **Materiale del corso:** sulla mia pagina web e su Teams



Nell'a.a. 2022-2023, questo laboratorio era tenuto dal Prof. Dario Della Monica.

Il laboratorio di quest'anno rimane del tutto uguale all'anno scorso:

- contenuti del corso
- modalità di lezioni
- modalità di esame



E' un laboratorio: (36 ore)

- **interattività**
- poche lezioni frontali
- alcune lezioni dimostrative con l'uso del laptop (per il linguaggio R)
- **modalità:**
 - durante le ore di laboratorio (a parte le lezioni frontali e dimostrative) lavorate al **progetto**
 - il docente è qui a disposizione per chiarimenti
- **scenario ideale:** il progetto viene svolto interamente o per maggior parte durante le 36 ore di laboratorio



Quando?

- ogni Venerdì 8:30 - 11:30
- possibili cambiamenti per conciliare gli impegni accademici e istituzionali del prof. Montanari
 - controllate Teams

Dove?

- Aula C2

Registrazioni:

- Cercherò di caricare puntualmente le registrazioni di ogni lezione
- I link alle lezioni del Prof. Della Monica dell'a.a. 2020/2021 sono disponibili qui:

https://users.dimi.uniud.it/~dario.dellamonica/teaching/22_23_1sem_BDlab/edition_20_21/link_alle_lezioni.txt



Libro:

- Basi di dati Atzeni-Ceri-Fraternali-Paraboschi-Torlone
Sesta edizione McGraw-Hill
- anche altre edizioni vanno bene; es.: Basi di dati: Modelli e linguaggi di interrogazione
Atzeni-Ceri-Paraboschi-Torlone Terza edizione
- (attenti alla differente numerazione dei capitoli)

Altro materiale (slides, altri libri, software, ...) sarà disponibile su Teams o sulla mia pagina web.



Progettazione e implementazione di una basi di dati

Gruppi di 3 o 4 persone

- pochissime eccezioni
- tali casi eccezionali sono sottoposti a regole più restrittive (e.g., scadenze più strette, etc...)

Votazione:

- voto unico
 - esame di teoria (Prof. Montanari): 75%
 - progetto di laboratorio (Dr. Geatti): 25%
 - incremento di 1 punto per i gruppi che consegnano in tempo (entro i primi 2 appelli)
 - è necessario raggiungere almeno la sufficienza (18) sia nella teoria che nel laboratorio



Formate i gruppi già a partire dai prossimi giorni

- scrivetemi una mail
- **UNA** mail per ogni gruppo (nè più nè meno)
- inviata da un membro del gruppo mettendo in **CC** tutti i membri del gruppo
- contenente (\forall membro del gruppo):
 - Nome e Cognome
 - numero di matricola
 - corso di laurea

Importante: tutte le comunicazioni riguardanti il progetto, devono includere in CC tutti i membri del gruppo.



Assegnazione dei progetti ai gruppi casuale, basata su politica First-In-First-Served

Potete proporre progetti “*alternativi*”:

- e.g., progetti su domini a scelta
- vanno prima di tutto discussi con me ed il Prof. Montanari



Tempistiche:

- Tempistiche flessibili per la consegna del progetto,
- l'esame di teoria viene conservato in attesa che il progetto venga ultimato e viceversa.
- **Suggerimento:** svolgere il progetto di pari passo con l'andamento del corso

Materiale da inviare:

- Relazione (da consegnare a me via email in formato pdf)
- Il focus deve cadere sulle fasi di progettazione (analisi dei requisiti, progettazione concettuale, logica e fisica) e non sull'implementazione (importanza ridotta in questo corso)
- Lunghezza: non c'è una regola ma tipicamente tra le 15 e le 30 pagine
 - scrivere tanto non significa aver fatto un buon progetto
 - scrivete poco ma bene



Valutazione

- volta anche a valutare contributi individuali
- interazioni dirette durante le lezioni di laboratorio
- dopo la consegna riceverete una valutazione (possono passare **VARIE SETTIMANE** dalla consegna alla valutazione):
 - a potete accettare la valutazione
 - b oppure potete scegliere di sostenere una discussione orale basata sulla relazione e sull'implementazione della base di dati



Importanza della corretta interazione tra le varie fasi della progettazione

Un errore nelle fasi iniziali del ciclo di produzione del software a conseguenze molto più gravi (difficoltà nella sua identificazione, risorse spese per correggerlo, propagazione alle fasi successive)

- Ogni fase produce uno o più documenti che rappresentano l'input per la fase successiva
- Coerenza e uniformità tra documenti prodotti nelle diverse fasi



Obiettivi del progetto - 2

- Prodotti della progettazione: non solo SQL ma anche e **SOPRATTUTTO** gli schemi prodotti dalle varie fasi (es. schema concettuale, logico, ecc.). Importanti perché:
 - rappresentano l'unica documentazione su cui basare la fase successiva
 - sono fondamentali (a volte l'unico supporto) all'implementazione (es., per l'implementazione di query)
- necessità di cicli tra fasi di progetto per raffinare/modificare prodotti di una fase precedente mentre si è in una fase successiva



Dimostrare di avere ben chiare le difficoltà e gli aspetti concettualmente critici ed interessanti della progettazione (ridondanze, generalizzazioni, vincoli aziendali, ecc.)

Warning:

- NO: aggiunte tecniche e complicate o estensioni troppo grandi del dominio. Ad esempio:
 - NON aggiungere 4 attributi derivati: 1 o 2 significativi bastano per mostrare che sapete eseguire l'analisi delle ridondanze
 - se si parla della gestione di aeroporti e compagnie aeree NON aggiungere (se non è nelle specifiche) tutta una parte nuova di dominio che riguarda, ad esempio, la gestione della vendita dei biglietti ai clienti



SI: aggiunte semplici di elementi mancanti importanti per analisi successive. Ad esempio:

- es. aggiunta obbligatoria: attributo derivato se non è presente nelle specifiche (con relative operazioni di lettura e modifica) per poter eseguire poi l'analisi delle ridondanze sull'attributo derivato aggiunto
- es. aggiunta obbligatoria: lista di operazioni frequenti con frequenze (7 o 8 son sufficienti, ma che permettano l'analisi delle ridondanze)
- esempio di aggiunta facoltativa: aggiungere uno storico ad una relazione (es.: tra impiegato e datore di lavoro) se lo schema concettuale appare troppo semplice



- 1 Analisi dei requisiti (dati e operazioni)
- 2 Progettazione concettuale (dati)
 - tiene conto del modello concettuale dei dati, totalmente astratto, astrae da modello logico e fisico (cosa)
 - produce uno schema concettuale
 - **diagramma E/R (o schema E/R)**



- 3 Progettazione logica (dati, operazioni usate per ottimizzazioni)
- tiene conto del modello logico dei dati, più concreto perché legato alla famiglia di DBMS (nel nostro caso, DBMS relazionali), astrae da modello fisico (cosa/come)
 - produce uno schema logico (noi adottiamo il paradigma relazionale)
 - ristrutturazione del modello E/R (analisi delle ridondanze basate su tavole dei valori e delle operazioni, eliminazione delle generalizzazioni, ecc. — di tutti i costrutti dell'E/R non esprimibili nel relazionale)
 - traduzione nello **schema relazionale**



- ④ Progettazione fisica (dati, operazioni usate per ottimizzazioni)
 - tiene conto di come vengono memorizzati fisicamente i dati (come)
 - produce il codice per la definizione delle relazioni ma soprattutto per la definizione degli indici
 - analisi e scelta di opportuni indici
 - definizione di relazioni e indici in SQL



- 5 implementazione (operazioni)
 - definizione di pochi (2 o 3) trigger risultanti dalla progettazione
 - definizione di poche (2 o 3) query significative (operazioni frequenti)
 - popolamento della base di dati (a mano o automatizzato, es.: R)

- 6 analisi dei dati in R
 - 2 o 3 esempi di query significative per una semplice analisi statistica (es.: trend o distribuzione di popolazione) realizzate interfacciando R al DBMS e visualizzazione del prodotto del risultato delle query attraverso opportuni grafic