Data Management and Database Design

Data Architecture Management (DAMG)

Week #1

Northeastern University

Course details

- Canvas
  - Pages

- Q&A Interaction
  - Slack

Course Overview

Data Management and Database Design

Course Information

DAMG6210 - Data Management and Database Design

Studies design of information systems from a data perspective for engineering and business applications; data modeling, including entity-relationship (E-R) and object approaches; user-centric information requirements and data sharing; fundamental concepts of database management systems (DBMS) and their applications; alternative data models, with emphasis on relational design; SQL; data normalization; data-driven application design for personal computer, server-based, enterprisewide, and Internet databases; and distributed data applications.

Next
Data Mgt and Database Design →

Assignment submissions

- Late submission leads to less marks
  - Each day late submission will cut 5 marks

  - If you fall behind assignments –
    - Difficult to catch up

- Waiting till last minute to deadline…
  - Will put you under stress and pressure

- I don't like to hear any excuses

Grading

- Course grades will be based on
  - participation, assignments, exams and presentations.

- Mid-term and Final exam must be taken at the time of the schedule

- Students are responsible for their own computing environments

- All students must agree to comply by NEU Academic Integrity Policy
  - http://www.northeastern.edu/osccr/academic-integrity-policy/

Attendance

- Attend lectures **regularly**

- **Taking Attendance is** randomly in nature

Topics for this course
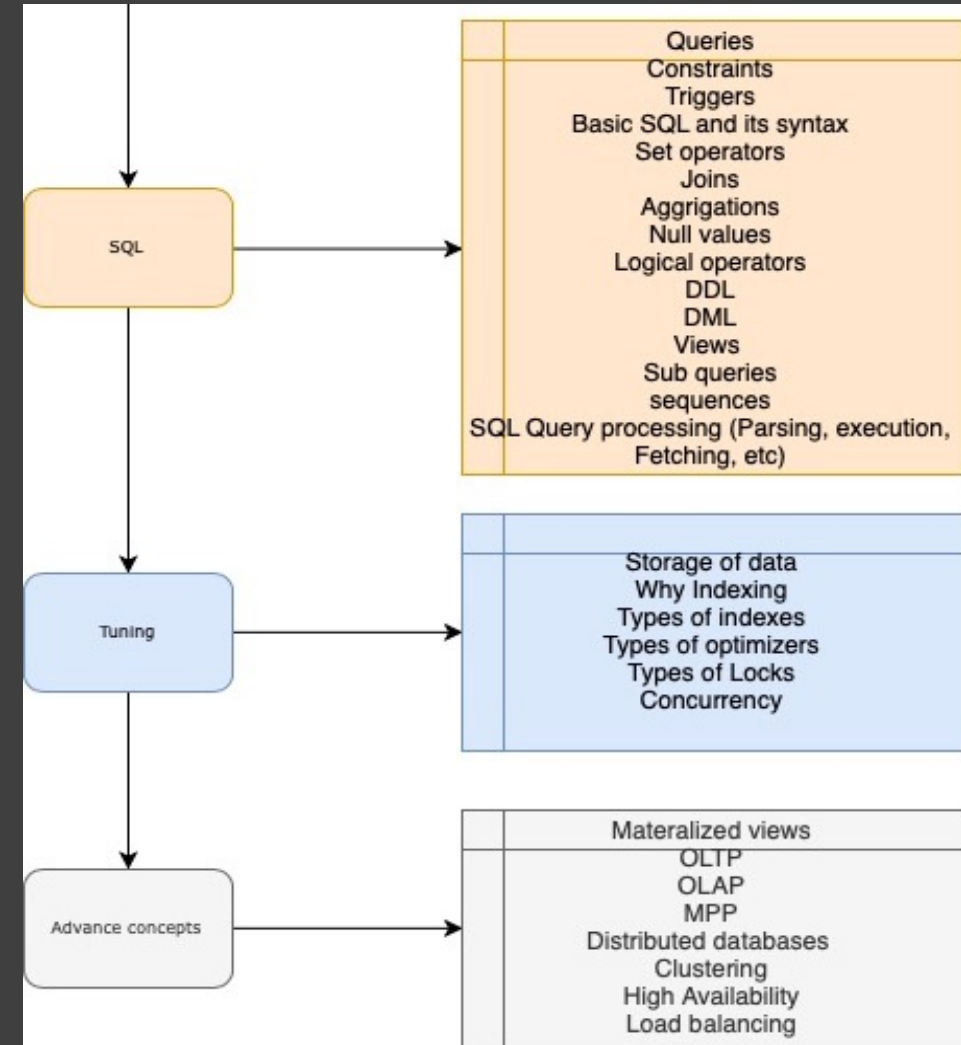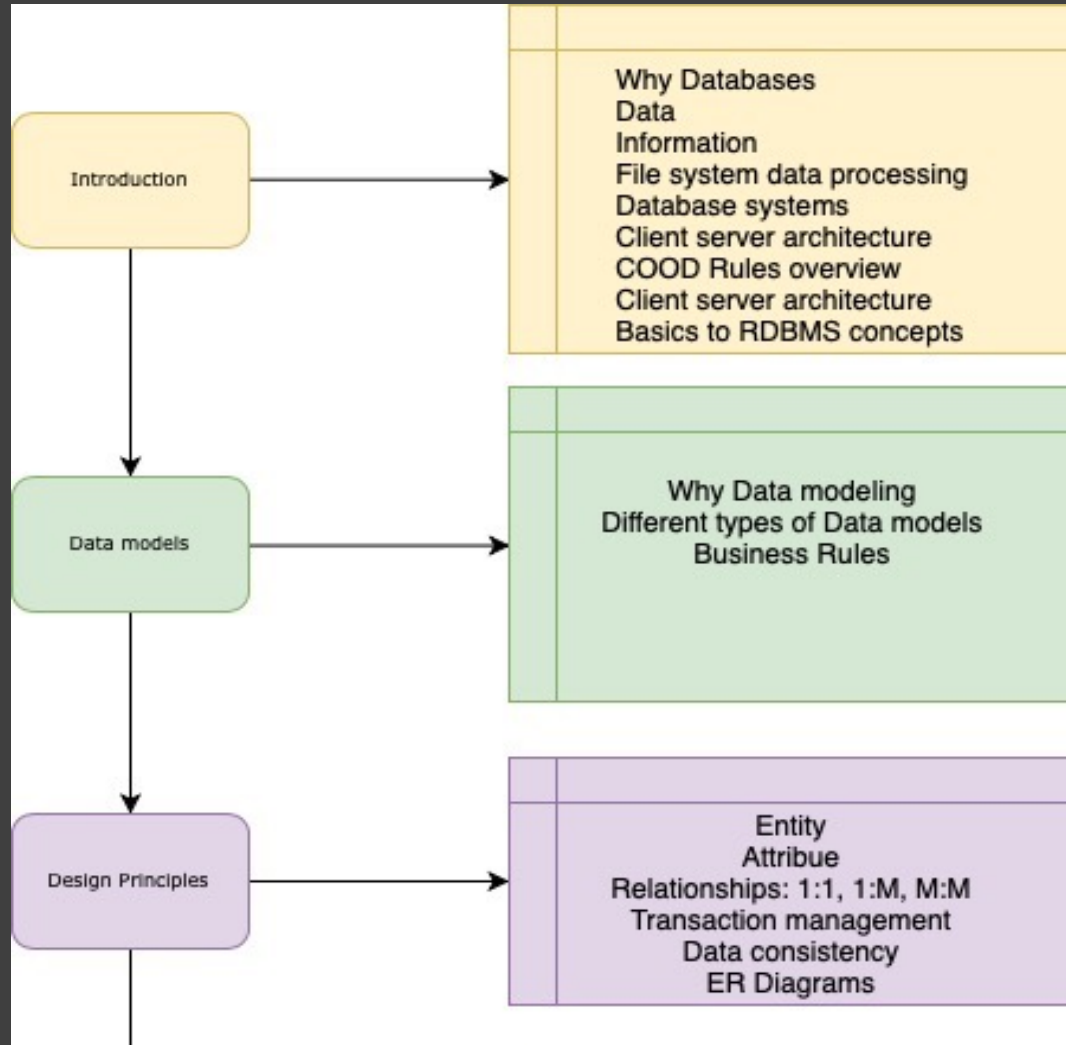
Presentations / Material

- These slides are for Lecture purpose only

- Will be uploaded to canvas after each class

Today's session

- Environments
- Why Databases
- Overview of –
  - Data
  - Information
  - File system data processing
  - Database systems
  - What are SQL databases
 - Why No-SQL Databases

Environment

- Different types of Environments
  - why we need them

- Source code management

- SDLC

- Virtual environment nomenclature
  - Host OS / Machine
  - Guest OS / Machine

- Software to create virtual machines
  - VM Ware
  - Parallels
  - Virtual Box
  - Hyper-V

Environment

- Different types of Environments
  - why we need them

- Source code management
- SDLC
- Virtual environment nomenclature
  - Host OS / Machine
  - Guest OS / Machine

- Different types of software to create virtual machines
  - VM Ware
  - Parallels
  - Virtual Box
  - Hyper-V

- For this course we will be using Oracle Cloud database
  - Introduction to cloud
  - On Premise vs Cloud
  - Advantages
  - Access for wherever you want and whenever you need.
  - Runs 24/7
  - Easily share the database
  - Collaboration
  - Security

Databases

What is Data?

What is Information?

Why we need Database system?

Why Data Management?
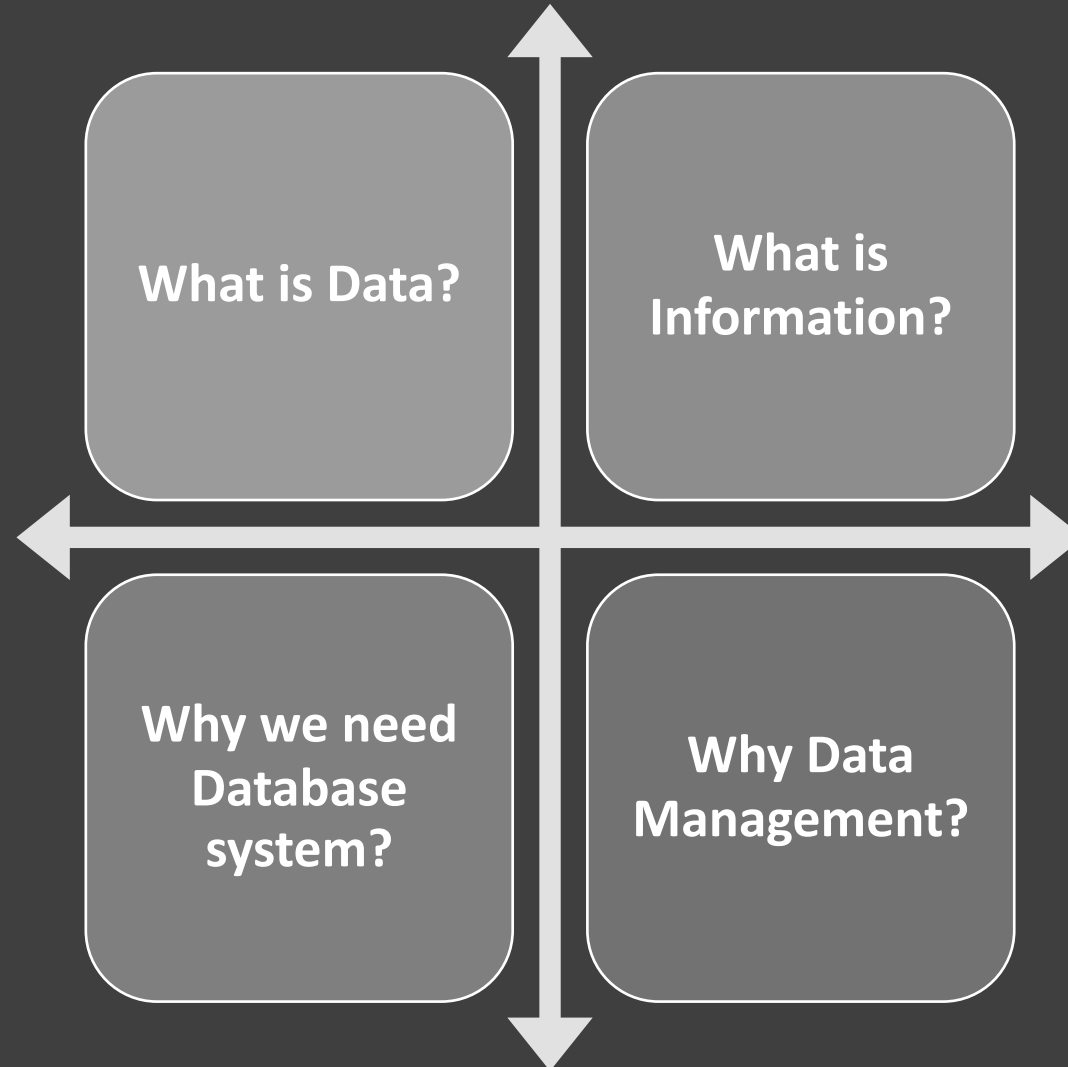
What is Data?

- Raw facts
  - data means a fact, an event
- Observations
  - just a cluster of numbers, words, or images
- Useful and non useful content
  - Data in itself does not really mean anything
- Not Organized
- Un Processed
- It is just content
  - Text
  - Numbers
  - Images, etc
- No dependency on Information

Which of the following describes you best?

- Student
- Employed
- Self employed
- Retired
- Not Employed

Gender:

- Male
- Female

Tell us how important of the following was in making your decision to use our airline:

	Very important	Important	Neutral	Unimportant	Not at all important
Departure and / or arrival time was more convenient	○	○	○	○	○
Flight had fewer stops or better connections	○	○	○	○	○
Air fare was better	○	○	○	○	○
Frequent flyer program	○	○	○	○	○
Inflight services are better (meals, movies, flight attendants, etc.)	○	○	○	○	○
Ground services are better (ticketing, baggage handling, check-in, etc.)	○	○	○	○	○
Personal preference for	○	○	○	○	○
Travel agent / company travel department recommendation	○	○	○	○	○
Aircraft preference	○	○	○	○	○

What is Information?

- Information is produced by processing Data

- It reveal's the meaning – Provides specifics and identify patterns

- Information cannot be processed without DATA

Raw data must be properly *formatted* for storage, processing, and presentation

Why we need Database systems?

- **Store**
- **Process**
- **Manage Data**
- **Solve problems**
  - What kind of problem?
    - Varity of data (Text, Images, Numbers)
    - Volume
    - Maintenance (Backup, Restore, Modifications)
    - Characteristics (Type, validation, relation)

So, Efficient way of data MANAGEMENT requires a *DATABASE*

Questions