

DA505 – Introduction to Data Modeling & Processing

Meeting Times and Locations

Saturday 13:00-16:00

Monday 19:00-22:00

Karaköy Minerva Han – 3005

Instructor

Dr. Banu Yoba

banuyobas@sabanciuniv.edu

Course Description, Aim and Content

We are currently in an era of data deluge. In many areas and domains, data are generated at a phenomenal speed that we have never experienced before. Given the large amount of data, one fundamental scientific challenge is how to model the data in a way that computational tools can analyze the data efficiently and effectively, revealing insight and make predictions.

This course aims to provide an overview of data modelling concepts and theories as well as hands-on practice and experience. In this course, you will learn and practice data modelling techniques, including: three phases of modelling, normalization, NoSQL concepts, new generation databases: key-value, wide column/ column, document and graph databases. Transaction management is also covered as it is crucial since transactions are the main source of big data.

The class consists of lectures by the instructor, class discussions in this area (which will change gradually as the area evolves), course project, and hands-on practices in each class using Lucid Chart for E-R and UML, SQL and related DBMSs. The course project can be selected considering the end of term project subject, in this way the data modelling part of the term project would be completed at the end of this course.

Topics

Introduction to Data, and Data Modelling Concepts

Conceptual Modelling

E-R Model

How to convert business requirements to E-R Diagrams

Entities, Relationships, Identifiers, PKs, Cardinality, FKs

Relational Database Management Principles

Logical Modelling

Converting a conceptual model to logical model

Integrity constraints

Normalization

Physical Modelling

SQL practices

Transaction Management Concepts

Consistency issues

Databases for Decision Support

Data warehousing Concepts

Architectures

Distributed Database Concepts
Conceptual understanding of Big Data and NoSQL
New generation Databases (MongoDB, Cassandra,
Key-value Databases
Wide column/ Column Databases
Document Databases
Graph Databases

Grading

Contribution to in-class discussions 10%

Assignment/Course Project 20%

The purpose of the course project is for you to learn hands-on experience of solving data modelling problems. Working as a group is permitted, and a team can consist of 1 to 3 persons. All members in one team will get the same grade.

Midterm Examination 25%

Final Examination 45%

References, Text Books

There are no required textbooks. Students may find the following books useful:

Modern database management, Jeffrey A. Hoffer V Ramesh; Heikki Topi 2013
Available at Student Reserve Collection QA76.9.D3 H64 2013

Veri tabanı sistemleri, Ünal Yarım an, 2010,
Call No: QA76.9.D3 Y37 2010

A relational model of data for large shared data banks, E. F. Codd, IBM Research Lab, San Jose, CA, Communications of the ACM, Volume 13 Issue 6, June 1970, Pages 377-387
Freely available at: <https://www.seas.upenn.edu/~zives/03f/cis550/codd.pdf>

Reading materials: electronic version of the required reading material will be available on course websites. The set of papers will change over the time with the development of the research area.

Announcements and SUCourse

A class website will be setup on SUCourse containing information about the course: syllabus, laboratory handouts, grades, miscellaneous information about weekly class activities, solution to the homework sets, discussion forums and an email directory of all people in the class. The lecture notes will be available on SUCourse for each class.

Students should follow the SUCourse site for this class regularly as they are responsible for all announcements and postings on this site.

Academic Integrity

All work submitted for a grade (including quizzes, home works and examinations) must be produced solely by the individual student/project team submitting it.

Any sources of information used in completing your work must be identified.
You are strongly advised to go through the academic integrity policy implemented in Sabancı University.

This document may be modified during the next semesters due to unforeseen reasons.