



Course guide

Databases and database design, 7.5 hp
D, final year – LP2 2011

Databases and database design, 7.5 hp

Welcome to the course databases and database design. In this course guide you find the syllabus, literature, schedule and description of learning activities we have and how they are examined – this guide shows in other words the content, how the activities fits in the schedule and what you can expect to learn.

The course builds upon that you as a student are active when you taking part of the content. Therefore, read carefully through this course guide so you know what we expect, the contribution from you. To maximize the outcome from the activities for learning it is important that you are well prepared for them, for instance the lectures and laboratory works. Through a positive and constructive approach you contribute to a stimulating and interesting time for you, peers and us as teachers. If something is unclear about the content in this course, it is a good idea to try to solve this matter as soon as possible. You are welcome to ask and suggest around the content as for example interests of learning, learning activities and others.

Teachers and contact information

Jesper Hakeröd, Michel Thomsen och Susanne Lindberg

Mail: Jesper.hakerod@hh.se; Michel.thomsen@hh.se; Susanne.lindberg@hh.se

Room: E408, E418, E411

Phone: 16 72 07, 16 75 03, 16 76 54

1. Course syllabus

Course objectives

The goals of the course are that the student should develop knowledge with regards to the principles and methods for the design and use of relational database systems. An additional aim is that the student should learn how to evaluate how alternative design solutions affect the functionality of relational databases. Through the course the student should develop an understanding of database concepts and acquire skills in solving basic problems concerning data storage in relational database systems.

Upon successful completion of the course the student should be able to:

Knowledge and understanding

- state the theories and basic concepts of the database area and reflect on the same
- explain the importance of normalization at analysis and verification of a database design solution

Skills and ability

- apply basic principles for relational databases in accordance with the relational model
- use object-orientated modeling in addition to graphic descriptive techniques as a tool for the design of relational databases which are the equivalent at the lowest to the third normal form
- manipulate (searching, adding, changing and deleting) of data and the modification and deletion of database objects (tables, keys, indexes, queries) at a basic level with the aid of the standardized database language SQL

Judgement and approach

- use normalization sufficiently well to be able to analyze and verify (together with modifying if the need arises) that a design solution fulfills at the lowest the third normal form
- relate to the Personal Data Protection Act (abbreviation PuL in Swedish) and be able to present what considerations have been made in terms of database design in order not to be in conflict with this Act

Primary contents

The course introduces database concepts (advantages and disadvantages, different user roles such as database administrator and database designer etc) together with basic design principles for relational design systems. The course includes a methodology for the development of database systems. It also provides the student with knowledge concerning why and how databases are used in companies and public authorities etc.

Teaching formats

Teaching consists of lectures, laboratory work and seminars. Lectures are interwoven with seminars and practical sessions to provide the student with both theoretical and practical knowledge of databases and database design.

Examination

The overall grades of Fail, Pass or Pass with distinction. The course has two examination components: the first part is a written examination (4.5 credits) based on the course literature, while the second part is based on seminar and practical session reports (3 credits).

Course Evaluation

Course evaluation is part of the course. This evaluation should offer guidance in the future development and planning of the course. Course evaluations should be documented and made available to the students.

Kurslitteratur

Conolly, T. & Begg, C. (2009). *Database Systems: A practical approach to design, implementation and management*. (5th edition), Addison Wesley, Harlow, England. ISBN 0-321-52306-7

Seminar and laboratory materials which is distributed in connection with the course.

Scientific papers available through the University library.

2. Schedule

Learning activities (preliminary – always check for changes in the schedule on the web)

Week	Date	Time	Activity
44	1/11	13-17	Course introduction, databases, environment and database architecture. Exercises is handed out. (JH)
	3/11	13-16	Relational model, normalization and E/R modeling part 1 (JH)
45	7/11	13-16	Relational model, normalization and E/R modeling part 2 (JH)
	9/11	13:00	Deadline for design exercises
	9/11	13-15	Supervision – design exercises (JH)
46	15/11	13-16	SQL (JH)
	16/11	13-15	Seminar for design and modeling (JH)
47	22/11	13-16	Conceptual-, logic- and physical design (JH)
	25/11	13-16	Laboratory work 1 – design of database (SL)
48	29/11	13-16	Database security, legal- and ethical aspects, transactions (JH)
	?	?	Laboratory work 2 – data views and transactions (SL)
	2/12	15:00	Deadline individuell uppgift – PUL
49	8/12	13-16	Theme: Databases in the web context (JH)
	9/12	08-10	Laboratory work 3 – Response times and denormalization (JH)
50	12/12	15:00	Deadline seminar & laboratory work
	13/12	13-16	Conclusions, questions and preparation for examination. (JH)

Examination

Week	Date	Examination	interest of learning (look in syllabus above)
49	9/12	Laboratory work and seminar	<ul style="list-style-type: none"> • apply basic principles for relational databases in accordance with the relational model • use object-orientated modeling in addition to graphic descriptive techniques as a tool for the design of relational databases which are the equivalent at the lowest to the third normal form • manipulate (searching, adding, changing and deleting) of data and the modification and deletion of database objects (tables, keys, indexes, queries) at a basic level with the aid of the standardized database language SQL • use normalization sufficiently well to be able to analyze and verify (together with modifying if the need arises) that a design solution fulfills at the lowest the third normal form • relate to the Personal Data Protection Act (abbreviation PuL in Swedish) and be able to present what considerations have been made in terms of database design in order not to be in conflict with this Act
51	?	Written examination	<ul style="list-style-type: none"> • state the theories and basic concepts of the database area and reflect on the same • explain the importance of normalization at analysis and verification of a database design solution • apply basic principles for relational databases in accordance with the relational model • use object-orientated modeling in addition to graphic descriptive techniques as a tool for the design of relational databases which are the equivalent at the lowest to the third normal form • manipulate (searching, adding, changing and deleting) of data and the modification and deletion of database objects (tables, keys, indexes, queries) at a basic level with the aid of the standardized database language SQL • use normalization sufficiently well to be able to

			analyze and verify (together with modifying if the need arises) that a design solution fulfills at the lowest the third normal form
--	--	--	---

3. Kursupplägg

To describe the course it can be divided into two parts, one part that aims to focusing on theories in the database field and one part aims to give you practical experience by designing a database. . It is important to be active in the learning activities to learn as much as possible.

The theoretical part builds upon lectures and the course literature. You as students have the possibility to learn the theory in the database field such as the relational model and normalization.

The practical part consists of exercises, seminars and laboratory work. You learn by doing things, reflect and test new things. You have peer reviews as a part of lectures the nextcoming week after each laboratory work. In the end this parts ends up in a examination of the seminars and laboratory works by a written report. The process for this is explained further for each learning activity below (and the weekly instructions for each activity). You work two and two with this part to get a better understanding by communication and by explaining for each other. Together you write the report (seminar and laboratory work).

After follows more detailed description of all the activities and the expected outcome.

As you can see it is in Swedish right now, but I put up a fully translated version until Thursday.

Design exercise and supervision

To get started to "thinking" databases and database design, it should first be good with some practice problems to solve in groups of two students. The exercises you should do is to practice and not examined, but is designed to prepare you for the design workshop and laboratory. Your solution is submitted via a website, whose address will be announced later. You will receive a proposed solution in return by mail. You will now compare the solution proposed by your submissions. Exercise and analyze what distinguishes the different solutions for and write down the lessons you drew from the analysis.

During supervision, you will get your account lessons for two other groups to thereby exchange experiences from the exercises. If you have questions from exercise, is the supervisor on site who can help to answer them.

Design Exercises (Exercises will be awarded in connection with the start of the course):

1. First Create an E/R model using ready-made relations, you stand for relationships and attributes.
2. Second Normalize a given table from unnormalized form to third normal form.
3. Third Design an E/R model from a case (text), validate it with normalization rules.

A checklist of what should be sent in your PDF:

- Cover page with all team members' name, surname and personal identity
- Describe any proposed solution on each side

Submit information via the designated website where you fill in your email address, upload the file and thus receive a solution proposal in return. Name your PDF file as follows:
exercise1_firstname1_lastname1_firstname2_lastname2.pdf

Translatione continues later during the day /Jesper

Vid behov av att komplettera designövningar

Om en grupp inte deltagit i övningen får gruppen komplettera genom att ladda upp filen för att erhålla lösningsförslaget som de ska jämföra med sin egen lösning. Gruppen ska även analysera minst en annan grupps lösningsförslag för att på så vis få motsvarande kunskaper. Kompletteringen redovisas skriftligen i en pdf-fil som skickas in via mejl till:

jesper.hakerod@hh.se; susanne.lindberg@hh.se; michel.thomsen@hh.se

Rubrik på ert mail ska vara [databaser & databasdesign – komplettering av designövning].

Designseminarium 1

Designseminarium 1 syftar till att lära sig hur vi kan skapa konceptuella och logiska modeller av en databas. Se förklarande information i häftet "Modellering mm" samt tillvägagångssätt i appendix D i kurslitteraturen.

En konceptuell modell visar på ett bra sätt de ursprungliga sambanden i och med att det går utmärkt att modellera samband av typerna 1:1, 1:* samt *:*. Med andra ord blir det lättare att förstå hur modellen avspeglar verkligheten. Ibland kan det vara svårt att bedöma en konceptuell modell om notationstekniken, tex UML, inte visar relationsobjekten som i sin tur kan innehålla attribut. För att vi ska kunna bedöma er modell utan några som helst tveksamheter behöver ni utveckla er konceptuella modell till en så kallad logisk design där relationsobjekten och deras attribut finns beskrivna (se exempel

i kurslitteraturen, figur 17.9). Ni ska använda er av notationen UML på samma vis som den beskrivs och används i kurslitteraturen, se bland annat kapitel 12.

Designseminariet genomförs i grupper om två studenter och är en delmängd av den slutliga inlämningen för "seminarium och laborationer" som lämnas in i slutet av kursen (ni lämnar inte in något förrän dess).

Inför designseminarium 1 ska ni ha:

- Läst igenom häftet "Modellering, mm"
- Läst igenom det aktuella caset som ni ska arbeta med och utarbetat ett förslag till lösning som presenteras med UML-notation (se kurslitteratur figur 17.9)
- Utarbetat ett data dictionary för samtliga attribut, tabell för tabell (se kurslitteratur figur 16.4). För val av datatyper hänvisas till:
http://www.teratrax.com/sql_guide/data_types/sql_server_data_types.html
- Förbered gärna frågor om ni undrar över något

Följande resultat bör vara uppnått efter designseminarium 1:

- En logisk E/R-modell gjord med UML-notation ska vara utvecklad utifrån den uppgift ni fått ovan (jämför figur 17.9 i kurslitteratur och 1c i häftet "Modellering mm"). För varje samband ska det finnas text som gör modellen tydlig och lätt att förstå. Alla benämningar/namn ska vara begripliga och nödvändiga nyckelattribut ska vara markerade.
- En metadatabeskrivning, så kallad data dictionary, som stämmer exakt mot er E/R-modell (jämför figur 16.4 i kurslitteraturen).
- Antaganden och andra överväganden för er modell redovisas och motiveras. Utgå ifrån vad som faktiskt beskrivs i uppgiften och undvik att hitta på massor av "vore bra att ha"-saker som inte finns med i er uppgift

Ni förväntas ta med er underlaget i samband med nästa föreläsningsspass för att i smågrupper utvärdera varandras lösningsförslag och ge konstruktiv kritik på vad som skulle kunna förbättras. Om ni har frågor finns handledare på plats för att hjälpa er besvara eventuella frågor.

Laboration 1

Ni fortsätter att arbeta med ert CASE. Ni ska utifrån er E/R-modell bygga databasen och mata in testdata, se anvisningar för laboration 1.

Ni förväntas efter laboration 1 ha:

- ER-modell i SQL Server. Observera att det ska finnas förtydligande text för varje samband i ER-modellen.

- Antaganden och andra överväganden för er modell redovisas och motiveras (i anslutning till ER-modellen).
- Data dictionary ska finnas för entiteter, attribut och samband.
- Ett schema över all SQL-kod för skapande av samtliga tabeller och samband i databasen. Observera att ni inte ska ta med INSERT-satser för data i schemat.
- Tre förslag till relevanta frågor som verksamheten kan tänkas behöva fråga databasen om.

Ni förväntas ta med er underlaget i samband med nästa föreläsningsspass för att i smågrupper utvärdera varandras lösningsförslag och ge konstruktiv kritik på vad som skulle kunna förbättras. Om ni har frågor finns handledare på plats för att hjälpa er besvara eventuella frågor.

Laboration 2

Under laboration 2 arbetar vi huvudsakligen med att läsa, skriva, uppdatera och ta bort data i databasen med hjälp av SQL för att tillgodose verksamhetens informationsbehov, se anvisningar för laboration 2. Bland annat behandlas stored procedures, triggers och vyer.

Ni förväntas efter laboration 2 ha:

- 1 uppdateringsexempel (UPDATE).
- 1 borttagningsexempel (DELETE).
- 1 insättningsexempel (INSERT).
- 3 informationsutdragsexempel (SELECT).
- 3 avancerade informationsutdragsexempel som hämtar information ifrån flera olika tabeller (SELECT).
- 3 frågor där ni använder er av aggregat eller andra funktioner, t.ex. SUM eller COUNT, AVG, MAX (se Connolly & Begg s. 149-151).
- 1 fråga som är nestad (subquery).
- 1 vy (VIEW).
- 1 stored procedure som uppdaterar kundtabellen med en ny kund (kunddata som in-parameter) och skapa en trigger som utlöses av samma stored procedure och skriver ut att en ny kund lagts till i tabellen.

Ni förväntas ta med er underlaget i samband med nästa föreläsningsspass för att i smågrupper utvärdera varandras lösningsförslag och ge konstruktiv kritik på vad som skulle kunna förbättras. Om ni har frågor finns handledare på plats för att hjälpa er besvara eventuella frågor.

Laboration 3

Under laboration 3 kopplar ni ihop ett färdigt webbgränssnitt med en relationsdatabas. Det finns även en databas med uppåt en miljon poster för att testköra frågor för att väcka tankar om hur lång svarstid vi får när vi ställer frågor, t ex om vi behöver denormalisera och använda kalkylerad redundans.

Ni förväntas ta med er era erfarenheter ifrån denna laboration i samband med avslutande föreläsningsspass för att i smågrupper diskutera varandras erfarenheter och tankar. Ni förväntas kunna muntligt kunna förklara vad ni tror är viktigast att tänka på utifrån ert genomförande av laboration 3.

Individuell uppgift – Personuppgiftslagen (PUL)

Syftet med denna uppgift är att medvetandegöra lagstiftning som påverkar innehållet i era databassystem. Det är inte allt som får eller ska lagras i en databas medan annan information kanske kräver samtycke för att lagras. På kurswebbsidan finns det 2 länkar till material som behandlar PUL att utgå ifrån tillsammans med en gästföreläsning av Anna Fredriksen. Försök därefter att reflektera över och förhålla dig till hur PUL påverkar lagring av data i databaser. Din reflektion förväntas omfatta 300-500 nedskrivna ord och ska tas med inför efterkommande diskussion i samband med nästa föreläsningstillfälle.

En checklista på vad som ska skickas in i er individuella uppgift (pdf-fil):

- Försättsblad med ditt förnamn, efternamn och personnummer
- omfattning 300-500 ords reflektion kring hur PUL påverkar lagring av data i databaser, du förväntas referera till anvisat kursmaterial och föreläsning i uppgiften

Skicka in uppgiften via anvisad webbsida där ni fyller i era mailadress, laddar upp filen och därigenom får ett lösningsförslag i retur. Döp er pdf-fil enligt följande:

PUL_förnamn_efternamn.pdf

I samband med efterföljande föreläsningstillfälle ska du ta tillsammans med 2 andra studenter jämföra och diskutera era reflektioner. Era slutsatser av jämförelsen och diskussionen ska resultera i att ni tillsammans lyfter fram de absolut viktigaste tankarna som ni direkt efteråt förväntas kunna redovisa muntligt.

Vid behov av att komplettera individuell uppgift - Personuppgiftslagen

Om du inte gjort reflektionen eller deltagit i efterföljande diskussionen behöver du ladda upp din individuella uppgift via anvisad webbsida. Du ska utöver din egen reflektion även jämföra den med en

annan grupps reflektion skriftligen för att på så vis få motsvarande kunskaper. Denna jämförelse redovisar du skriftligen i en pdf-fil som skickas in via mejl till:

jesper.hakerod@hh.se; susanne.lindberg@hh.se; michel.thomsen@hh.se

Rubrik på ert mail ska vara [databaser & databasdesign – komplettering av individuell uppgift - PUL].

Examination av seminarium och laborationer

Designseminarium, laboration 1 och laboration 2 samt det avslutande seminariet ligger till grund för den slutliga inlämningen av seminarium och laborationer.

Inlämningen ska innehålla:

- Försättsblad med namn, efternamn, personnummer och mailadress till samtliga medförfattare
- ER-diagram (utdrag från SQL Server) inklusive eventuella förklaringar
- Data Dictionary (enl. en beskrivning så att de tar med allt)
- Queries från labb 2:
 - UPDATE x1
 - DELETE x1
 - INSERT x1
 - SELECT x1
 - Avancerade SELECT x3
 - Aggregerande funktioner x3
 - Nestad query x1
 - VIEW x1
 - Stored procedure x1
 - trigger x1
- analys och verifiering av designlösningen i förhållande till uppfyllande av tredje normalform
- en kort uppsummering av de erfarenheter ni fick under laboration 3

De frågor som presenteras ska vara genomtänkta och ha ett tydligt syfte i företagets användning av databasen, t.ex. varför deras vy har ett visst innehåll, eller varför just den valda stored proceduren valdes.

De kriterier som vi speciellt är intresserade av i samband med bedömningen är:

- att grundläggande principer för relationsdatabaser i enlighet med den relationella modellen tillämpas korrekt
- att ni visar god kunskap och bredd i er användning av frågespråket SQL för att möta de behov som finns i CASET
- att ni har gjort en korrekt analys och verifiering av designlösningen i förhållande till uppfyllande av tredje normalform

Skicka in uppgiften via anvisad webbsida där ni fyller i era mailadress, laddar upp filen och därigenom får en bekräftelse av inskicket i retur. Döp er pdf-fil enligt följande:
uppgift1_förnamn1_efternamn1_förnamn2_efternamn2.pdf

If you need a supplementary examination of seminar and laboratory works

If you have not finished seminar and laboratory work, you can complete it afterwards by upload the pdf-file on the suggested webpage and send a mail with the same pdf-file attached to:

jesper.hakerod@hh.se;susanne.lindberg@hh.se;michel.thomsen@hh.se

The subject of your mail should be [databases and database design – supplement of seminars and laborations].

Lecture 1 – course introduction

A course introduction to present learning activities, interests of learning, literature and schedule of the course. Lecture 1 deals with databases, the environment and architecture. Literature: chapter 1-3 in Connolly & Begg.

Lecture 2

The lecture deals with the relational model, normalization and E/R-modeling (conceptual and logical design).

Literature: chapter 4, 12, 14 in Connolly & Begg.

Lecture 3

We continue with the relational model, normalization and E/R-modeling (conceptual and logical design).

Literature: chapter 4, 12, 14 in Connolly & Begg.

Lecture 4

The lecture deals with communication with the database by using SQL.

Literature: chapter 6-8 in Connolly & Begg.

Lecture 5

The lecture deals with conceptual- logical and physical database design.

Literature: chapter 16-18 in Connolly & Begg.

Lecture 6

The Lecture deals with security and database administration, legal and ethical aspects and transactions.

Literature: chapter 20- 22 in Connolly & Begg.

Lecture 7 Topic lecture about web applications

The Lecture deals with how web applications can use relational databases.

Literature: chapter 30-31 in Connolly & Begg.

Lecture 8 – kursavslutning

Finishing off the course, conclusions and coming examination. We answer the last questions that you might have and a walk through some questions in the written examination to prepare for it.

Reading instructions for Connolly & Begg (2010):

- chapter 1-3 (not 3.6)
- chapter 4, 6-7 (not 4.2.2)
- chapter 8 know about stored procedures and triggers is
- chapter 9 know what QBE is
- chapter 10-12, 14, 16-19
- chapter 20-22 (not 20.3-20.5, not 21.2 except to know about EU directive in 21.2.4, not 21.4, not 22.4-22.5)
- chapter 24, 30 (not 24.1.3, not 24.2, not 24.3.2-24.3.3, not 24.5, not 30.3-30.9)
- chapter 31 know what XML is
- chapter 32-33, 35 (not 32.5, not 33.7, not 35.6)

Good luck with the learning activities!

Jesper