



IV2014 Data Warehousing 7.5 credits

Data Warehousing

This is a translation of the Swedish, legally binding, course syllabus.

If the course is discontinued, students may request to be examined during the following two academic years

Establishment

Course syllabus for IV2014 valid from Autumn 2010

Grading scale

A, B, C, D, E, FX, F

Education cycle

Second cycle

Main field of study

Specific prerequisites

Eligibility for "single-course" students not enrolled in a KTH programme:

- BSc degree within Business Administration, Economics, 180 ECTS credits (hp) in Technology or Natural Sciences or equivalent and
- documented proficiency in English B or equivalent.

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

The general goal with this course is to familiarize the students with a special kind of information systems called Data Warehouses – their role, utilization and benefits for organisations, as well as architectures and underlying technologies relevant for their development.

After taking this course the student should have achieved the following objectives:

1. knows central terminology in the area, specifically the terminology introduced by Ralph Kimball
2. is able to produce and document dimensional models for a data warehouse based on an informal domain description
3. from a given source, produce routines for data transfer into a data warehouse
4. is able to implement dimensional models in a given system, populate these models with data, and use a front-end systems for extracting and analyzing the data present in a data warehouse. Can use a given ETL system to extract data from different files and load it into relational tables
5. can summarize, present and assess results from research literature in the area

Course contents

The course contains the following parts:

1. Information requirements and information modelling relevant for enterprise management level
2. History, state-of-the-art, trends for data warehouse
3. Multidimensional modelling as a practical method for Data Warehouse design
4. The ETL process
5. Project management and lifecycle during the development and the maintenance of a Data Warehouse
6. Decision support systems
7. Business intelligence
8. Data Mining

Disposition

Lectures, lessons and seminars.

Course literature

Preliminary:

- R. Kimball & M. Ross: The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, 2nd Edition (Edition: 2), John Wiley & Sons, Inc., 2002, 0-471-20024-7
- Compendium

Examination

- INL1 - Assignment, 3.0 credits, grading scale: P, F
- TEN1 - Examination, 4.5 credits, grading scale: A, B, C, D, E, FX, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

The examination for the course consists of two parts:

- Written exam
- Assignments

The assignments part contains four assignments.

- Assignment 1: Dimensional Modelling
- Assignment 2: ETL
- Assignment 3: Analyse Services
- Assignment 4: Article Presentation

The assignments shall be conducted in groups of 4 students.

GRADING CRITERIA FOR ACHIEVING THE OBJECTIVES

The grading criteria for achieving the different objectives of the course are presented below.

1. Knows central terminology in the area, specifically the terminology introduced by Ralph Kimball

A Is able to correctly define, exemplify, use and compare at least 95% from a sample of central concepts within data warehousing area, specifically from the data warehousing architecture area, dimensional modelling, the ETL process, and from-end tools

B Is able to correctly define, exemplify, use and compare at least 90% from a sample of central concepts within data warehousing area, specifically from the data warehousing architecture area, dimensional modelling, the ETL process, and from-end tools

C Is able to correctly define, exemplify, use and compare at least 85% from a sample of central concepts within data warehousing area, specifically from the data warehousing architecture area, dimensional modelling, the ETL process, and from-end tools

D Is able to correctly define, exemplify, use and compare at least 80% from a sample of central concepts within data warehousing area, specifically from the data warehousing architecture area, dimensional modelling, the ETL process, and from-end tools

E Is able to correctly define, exemplify, use and compare at least 75% from a sample of central concepts within data warehousing area, specifically from the data warehousing architecture area, dimensional modelling, the ETL process, and from-end tools

Fx Is able to correctly define, exemplify, use and compare at least 60% from a sample of central concepts within data warehousing area, specifically from the data warehousing architecture area, dimensional modelling, the ETL process, and from-end tools

2. Is able to produce and document dimensional models for a data warehouse based on an informal domain description

To get a certain grade the student is expected to be able to produce a model with the following properties.

A B C D E Fx

Syntactic correct Max one error Max one error Max one error Max one error Max one error
Max one error

Semantic correct 95% 90% 80% 70% 60% 50%

Complete 95% 90% 80% 70% 60% 50%

Not-redundant 5% 20% 30% 50% 75% 75%

Easily comprehensible Absolutely Mostly To large degree To large degree Acceptable Acceptable

Syntactic correct means that the language and/or the drawing notation are used correctly.

Semantic correct means that the content is correct.

Complete means that all important parts of the domain description are captured in the model.

Not-redundant means that none unnecessary details are included in the model.

3. From a number of given sources, can produce models and routines for data transfer into a data warehouse

As above (i.e. as the grading criteria for achieving objective 2)

4. Is able to implement dimensional models in a given system, populate these models with data, and use a front-end systems for extracting and analyzing the data present in a data warehouse. Can use a given ETL system to extract data from different files and load it into relational tables.

Passed: Implement predefined models in a given data warehousing system, populate these models with data available in the data staging area, use two front-end systems for extracting specific data which is needed for answering some predefined questions. Compare the two front-end systems. Use a given ETL system for extracting data from a number of sources and loading it into relational tables.

Fx: Partially implement predefined models in a given data warehousing system, populate these models with data available in the data staging area, partially use two front-end systems for extracting specific data which is needed for answering some predefined questions. Partially, use a given ETL system for extracting data from a number of sources and loading it to relational tables.

5. Can summarize, present and assess results from research literature in the area

Passed: Without teacher assistance, be able to read and understand the content of a selected article, as well as present and discuss the article.

Fx With some teacher assistance, be able to read and understand the content of a selected article, as well as present and discuss the article.

GRADING CRITERIA FOR EXAMINATION OF PARTS OF THE COURSE

The course is examined through a written exam (4.5 points) and a set of assignments (3 points).

Written Exam

The written exam assesses the achievement of objectives 1, 2 and 3. The following shall be fulfilled in order to get a certain grade on the written exam.

A All the objectives (i.e. 1, 2 and 3) shall be achieved with grade A

B All the objectives (i.e. 1, 2 and 3) shall be achieved at least with grade B

C All the objectives (i.e. 1, 2 and 3) shall be achieved at least with grade C

D All the objectives (i.e. 1, 2 and 3) shall be achieved at least with grade D

E All the objectives (i.e. 1, 2 and 3) shall be achieved at least with grade E

Fx All the objectives (i.e. 1, 2 and 3) shall be achieved at least with grade Fx

Assignments

Assignment 1 is used to assess the achievement of objective 2

Assignment 2 is used to assess the achievement of objective 3

Assignment 3 is used to assess the achievement of objective 4

Assignment 4 is used to assess the achievement of objective 5

The set of assignments is graded according to the following:

Passed: objectives 2 and 3 shall be achieved at least with grade C and objectives 4 and 5 shall be achieved with the grade Passed

Fx objectives 2, 3, 4 and 5 shall be achieved at least with grade Fx

Other requirements for final grade

In order to pass the course, the student needs to pass both the written exam and the assignments. Final grade is based on the grade of the written exam.

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.