Assignment 2 - IN5040-H2025

General Information

In this mandatory exercise you will learn about advanced database technologies. The questions you have to answer could be examination questions.

Delivery Instructions – Assignment 2

Where:	Devilry - https://devilry.ifi.uio.no
Who:	Alone
Due:	Hard Deadline: 16 October 2025 - 11:59 AM
Deliverable:	One PDF file that contains the answers to the theoretical questions.
Questions:	Vera Goebel (goebel@ifi.uio.no) Thomas Plagemann (plageman@ifi.uio.no)

Administrative Issues

You are not meant to copy/paste from the course material or the syllabus (pensum). Instead you should read, understand and write the answers in your own words. Only students that have passed both parts of the mandatory exercise can take part in the oral examination!

Question 1. Text – DSPS (~ 1 PAGE)

a) Explain the motivation for Data Stream Processing Systems (DSPS).
b) Elaborate on the differences between a traditional DBS and DSPS.
c) Which applications make use of DSPS?

Question 2. Text – DSPS (~ 0.5 PAGE)

a) What is a window in DSPS and what are windows used for?
b) What is the difference between tumbling and sliding windows and what are their advantages and disadvantages?

Question 3. Text (~ 0.5 PAGE) Complex Event Processing

a) What does the every and followed-by (->) patterns do?
b) Explain the difference between the following:
    a. every A -> B
    b. A -> every B

Question 4. Text – Distributed DBS (~1 PAGE)

a) Explain the difference between a centralized DBS and a distributed DBS.
b) Why do we need distributed DBS?
c) What are the advantages and disadvantages of distributed DBS?
d) How can you classify the different distributed DBS solutions?
(Hint: there are three characteristics/dimensions)
e) Elaborate on the different approaches to build distributed DBS.

Question 5. Text - Database Integration (~ 1 PAGE)

a) Explain the challenges and problems for Database Integration (MDBS).
b) What is a heterogenous Multi-database system (MDBS), and how does a MDBS work?
c) Explain the concepts of the integration layer and explain the issues related to integration.

Question 6. Text - Web Data Management (~ 1 PAGE)

a) Explain the problems and approaches for Web Search and Querying.
b) Explain the approaches for Web Data Exchange.
c) Explain the challenges and approaches for Web Data Integration.

Question 7. Text – ML Data Management & Data Warehousing (~ 1 PAGE)

a) Name and explain the challenges and tasks for ML Data Management.
b) Briefly describe the stages of the Knowledge Discovery in DBs process.
c) Briefly describe the Data Warehouse (DW) Life Cycle.
d) Name the different relational DW design schemas. Briefly explain the motivation for each schema, and any possible disadvantages it may have.

Question 8. Text – Scalable Data Management (~ 1 PAGE)

a) How can you design Scalable Data Management solutions for Big Data?
b) Explain the challenges/problems and approaches for cloud computing.
c) Explain the CAP Theorem and the different NoSQL solutions.