Examination Information

This folder contains multiple files, which is going to be filled by you with the solutions to each assignment. This file describes the tasks you must perform, while the other files are referred to via the tasks described below. Each task has a percentage assigned to it that gives you an indicator of how much the task weighs towards your grade for the examination. **Remember that all SQL submitted at this exam must run in PostgreSQL – other dialects are not acceptable**. Mind your time usage on each task (each task has a time recommendation attached) but remember there are multiple subtasks for each task! The following sections describe the tasks you need to perform. READ the entire document before solving the assignments!

Each task can have subtasks – remember to read them all before you start. The tasks present in this exam:

- Task 1
  - Subtask 1 – ER and EER diagrams
  - Subtask 2 – Mapping to Tables
  - Subtask 3 – Querying the Database
- Task 2
  - Subtask 1 - Normalization
- Task 3
  - Question 1 – Document Based Databases
  - Question 2 – Graph Based Databases
  - Question 3 – CAP theorem
- Task 4
  - RegEx 1 – Find the king and queen!
  - RegEx 2 – Find all the yelling!
- Submission

Task 1: Relational Databases – Part 1 (3 subtasks) (60% ~ 72 min)

Subtask 1: ER and EER diagrams (30 min)

This subtask concerns itself with generating an ER/EER diagram from a textual context. For this task, make sure you either open the Task1Subtask1.drawio file (if you have the offline version of draw.io), or download your diagram and override the Task1Subtask1.drawio file. If you are using other tools, make sure you export it into a commonly readable format such as PNG, JPG, or PDF, and call the file Task1Subtask1 when done. You are allowed to draw the diagram by hand and attach a picture of it instead but be very sure everything is easily readable as we cannot grade anything we cannot read.

The text below is the output of an interview. This is the only documentation you have, so if something is unclear, you will have to make the decision on how to understand it - just make a note in the diagram explaining your reasoning. Remember to be as precise as possible and include all the types of descriptions you have learned (some might not be applicable for the case though). This includes strong and weak entities, inheritance, relationships, cardinalities, keys and more.

Interview Results (in Danish):

> *Et ny opstartet firma, Omikatu Cloud Systems, ønsker at skabe et nyt produkt til at gemme, versionere og dele filer over nettet som en cloud tjeneste.*
>
> *Filerne skal ligge i en flad fil struktur i én mappe, hvor at alle filernes navne har et fortløbende nummer (1.docx, 2.xlxs, 3.jpg). Brugerne af systemet skal dog kunne oprette mappestrukturer i deres interface. En fil skal for brugeren se ud som om at den ligger deres egne oprettere mapper, inklusiv et sigende navn som brugeren fastsætter for filen selv (så fx /minmappe/mit_dokument.docx i databasen, og 2342.docx i filsystemet). Systemet skal kunne håndtere flere brugere af samme system, så de skal have en konto og skulle kunne logge på systemet med deres e-mail og kode (ignorer sikkerhedshensyn), og deres navn skal desuden også gemmes. Filer som brugerne uploader skal kunne overskrives af en nyere version, men den gamle gemmes så de altid kan finde den igen. Omikatu Cloud Systems ønsker desuden at en fil skal kunne deles med en anden bruger, men selve delingen skal have en udløbsdato. De delte filer kan ikke ændres af dem som filen er delt med.*
>
> *I systemet understøtter man ikke alle filtyper (med vilje), og ønsker at begrænse brugerne til at kun at kunne gemme bestemte typer alt efter hvilken licens type de har købt. Der er nuværende defineret 3 licens typer: Gratis, Pro, og Enterprise. Hver licens type begrænser også hvor mange fysiske computere der må have deres program installeret lokalt, samt typen af filer en bruger kan gemme. De forskellige licenser har følgende begrænsninger:*
>
> 1. *Gratis må tilslutte en Computer og må bruge .docx, og .txt*
> 2. *Pro må tilslutte 3 Computere og må bruge .docx, .txt, .pdf, .xlsx, .jpg, og .png*
> 3. *Enterprise må tilslutte alle de computere de vil, og bruge alle dokumenttyper*

Subtask 2: Mapping to Tables (30 min)

This subtask concerns itself with mapping an ER diagram (or EER) to tables in a database.

For this task, make sure you open the Task1Subtask2and3.sql file. The file is empty, and you will have to fill it out with the database create script for the ER diagram you made in ***Task1 Subtask 1.*** If you were not able to create the diagram, use the interview result text as the basis for your tables.

Subtask 3: Querying a database (12 min)

This subtask concerns itself with querying databases. Open Task1Subtask2and3.sql, and add the following queries at the end of the file:

1. Add two entries to each table.
2. Query for all users with the “Pro” license.
3. Query for all files shared with another person, and output the name of the file, the name of the owner of the file, and name of the user the file was shared with, and lastly, a timestamp for when the share runs out. I expect a result like this:
    a. | DataStore.docx | Jens Hansen | Anne Ludwigsen | 2022-05-16 10:32 |

Task 2: Relational Databases Part 2 - Normalization (20% ~ 24 min)

This task concerns itself with normalizing a given dataset (the excel image below) to tables in SQL, and making a query within the normalized tables

The output of this task should be saved in the Task2.txt file, which is currently empty. Open the Task2.txt file directly and save to that file using your editor. Feel free to write comments.

Absurdity Rollercosters Incorperated's Log Entries

Note	Approval	Log entry by	Average Wait time	Ride
21/04/2015 - Standard Operations		Jake Hanson	35 minutes	The Death Trap
01/05/2022 - Standard Operations		John Doe	2 hours	Roller Blades
Yesterday morning - Dont know what happened! Just did not work! Repair by Peter Johnson	Approved	Ann Coster		Mustard Mountain
First Quarter 2021 - Inspection of motors and rails, By Dan Johanson	Approved	Daniel the Park Ranger		Mustard Mountain
May 4, 2022- Restraint failed, customer got eaten by hippo. Customer could not be retrieved. Ann Coster responded.	Not Approved			Hungry Hungry Hippoes
2022-06-9 - Students got stuck in gears. One arm is still missing. Repair and cleaning by Peter Johnson	Approved	Ann Coster		The Death Trap
2022-06-09 - Standard Operations		Ann Coster	3 hours 5 min	Hungry Hungry Hippoes
Ride would not start. Arm found stuck in electric parts. Removed and cleared. Repair by Peter Johnson. 10th of may 2022	Approved	John Doe		The Death Trap

Tasks are as follows (read all steps first!):

- Normalize the dataset to at least the third normal form.
- Write the normalized schema in Task2.txt using the template/shorthand form (for example ***TableName(PK id, prop1, prop2, FK prop3)***)
    - Alternatively, use “create table” statements (your choice) if you are not comfortable with the template/shorthand notation.
- You do **NOT** have to insert the data into tables!
- **Remember** the relationships with primary and foreign keys – these are essential for this assignment!
- Feel free to write notes/reasoning for your choices in the document.

Task 3: Database choices (10% ~ 12 min)

Open the file called "Task3.txt" from the zip file. The questions below are also stated within the file. Answer all questions to the best of your ability, and **make sure to motivate your answer**! Your answers, together with your considerations, must be written in Task3.txt

1. When should one choose a document-based NoSQL database over a relational database?
2. When should one choose a graph-based database over a document-based database?
3. What is the CAP theorem, and how would you explain it?

Task 4: Regular Expressions (10% ~ 12 min)

Open the file called "Task4.txt" from the zip file. The file contains text under the line. You will have to create two regular expressions that ONLY match the stated assignment. Write the answers in the Task4.txt file in the top next to the questions in the top. The assignments are as follows:

1. Match the entire sentence when it contains the word Queen or King.
2. Match all words with only CAPITAL letters (STORE BOGSTAVER).

Submission

To submit the assignment, make sure you have a folder containing the following files:

- Task1Subtask1.drawio (alternatively .jpg, .png, or .pdf)
- Task1Subtask2and3.sql
- Task2.txt
- Task3.txt
- Task4.txt

Instructions:

1. Check each file has your work contained within it.
2. Zip the files without the folder containing them (mark all files and zip them together).
3. Name the zip file after your SDU Username such as "abcd17.zip.
4. Go to the assignment location from where you downloaded the counting activity.
5. Upload the zip file.
6. For your own security, download the file again and unzip it to verify the version stored on the server works as intended (and that you included the right files – yes this happens!).