# CSC7072: Databases, fall 2015

Dr. Kim Bauters

how to prepare for the DB exam

# Revision

## group project

*first things first:* group project

able to work together during lab,
as well as during Friday's and Monday's lecture (no more lectures!)

step-by-step conversion of the ER model into database schema, (3) any normalisation that has been applied and why/or why not, and (4) an explanation of the aforementioned queries. ***(5) In addition, the report should explain in some detail which student contributed to each part of the process.*** The project report should be no more than 15-20 pages

get in touch if at any time you feel that
the peer assessment is unfair to you

remember that the peer assessment for is only there to
evaluate the measurable contributions of each team member

# Revision

having questions?

starting to study?

if you have questions:

Tuesday 8th, 9 – 10am
Tuesday 8th, 11am – noon
Tuesday 8th, 3pm – 4.30pm

Tuesday 15th, 9 – 10am
Tuesday 15th, 11am – noon
Tuesday 15th, 3pm – 4.30pm

contact me in advance to arrange a slot to address your questions!

# Revision

## know what to expect

what does the exam look like?

you get 5 questions, you have to answer *any* 3 questions

1. SQL (+ relational model)
2. ER modelling
3. normalisation
4. data mining
5. transactions

don’t select them beforehand …

# Revision

## know what to expect

what does the exam look like?

1. SQL (+ relational model)

   you need to know all basic concepts from 1 and 2

   - what is a table? attribute? domain? tuple? relational schema?
   - what is the difference between a relation schema/instance?
   - what is a database? bad database? superkey? candidate key?
   - what is a declarative language? a foreign key constraint?
   - what is the input of a SQL query? the output?
   - what is a natural join? inner join? outer join?

# Revision

## know what to expect

what does the exam look like?

1. SQL (+ relational model)

   you need to know SQL, both DML and DDL – 03, 03a, 04, and 06

```
SELECT {attribute [AS new_attribute_name]}
    FROM {table [AS new_table_name]}
        [{JOIN table ON attribute = attribute}]
        [WHERE {condition}]
        [GROUP BY {attribute}]
            [HAVING {condition}]
        [ORDER BY {attribute}]
```

# Revision

## know what to expect

what does the exam look like?

1. SQL (+ relational model)

   you need to know SQL, both DML and DDL – 03, 03a, 04, and 06

   CREATE VIEW view_name AS sql_expression

   INSERT INTO table [({attribute})] VALUES ({value})

   UPDATE table SET {attribute = value} WHERE {condition}

   DELETE FROM table WHERE {condition}

# Revision

## know what to expect

what does the exam look like?

1. SQL (+ relational model)

   you need to know SQL, both DML and DDL – 03, 03a, 04, and 06

```
CREATE TABLE advisor (
    student_id        VARCHAR(5),
    instructor_id     VARCHAR(5),
    PRIMARY KEY (student_id),
    FOREIGN KEY (instructor_id) REFERENCES instructor (id)
        ON DELETE SET NULL,
    FOREIGN KEY (student_id) REFERENCES student (id)
        ON DELETE CASCADE  )
```

# Revision

## know what to expect

what does the exam look like?

1. SQL (+ relational model)

   you need to know SQL, both DML and DDL – 03, 03a, 04, and 06

   - what is a derived relation?
   - what is a subquery? correlated subquery?
   - what is a single row/multiple row/multiple column subquery? how and when do you use each one? (*e.g.* give example)
   - similarity/differences between derived relation/view?
   - and restrictions on queries: “create with a subquery”, “use EXISTS” …

# Revision

## know what to expect

what does the exam look like?

❷ ER modelling

just two things – 05, 07, and 07a:

- convert a problem into an ER model;
- convert that ER model into relational schemas.

time management will be critical! – don't forget 08a for scoring!

# Revision

know what to expect

what does the exam look like?

❸ normalisation

define, identify, transform

- what is normalisation? what is 0NF, 1NF, 2NF, 3NF, BCNF, 4NF, 5NF?
- what is the normal form of this relational schema?
- convert this schema into 1NF, 2NF, 3NF
- give examples of a schema not in a given NF

*basically:* everything!

# Revision

## know what to expect

what does the exam look like?

4. data mining

   what is it? why do we need it? do it yourself

   - what is data mining? what are data mining tasks?
   - what is the general approach to data mining?
   - what is Apriori? a transaction? Market Basket Analysis?
   - what is confidence? support? data model? item? transaction set?
   - what is an itemset? what is an association rules?
   - be able to apply Apriori algorithm on a given transaction set

*basically:* everything!

# Revision

## know what to expect

what does the exam look like?

5. transactions

   what is it? why do we need it? do it yourself

   - what is a transaction? ACID?
   - what is atomicity? consistency? isolation? durability?
   - what is permanent memory? volatile/non-volatile memory?
   - what are the stages of a transaction? what is a control scheme?
   - what is a conflict? conflict-serialisability? log-based recovery?
   - and be able to apply conflict-serialisability/log-based recovery

*basically:* everything!

# Revision

## know how to study

how do you study for this course?

1. SQL (+ relational model)

   use PeerWise for factual knowledge

   use PeerWise for common SQL errors

   lots and lots of exercises! (make all 50 queries)

*don't forget:* everyone who contributes at least 5 decent questions to PeerWise will get bonus points (+5% on assessments)

repeat often! ***(spaced repetition)***

memory loss in 1 min, 15 min, 1 hour, 1 day, 1 week, 1 month, 1 year, never

# Revision

## know how to study

how do you study for this course?

❷ ER modelling

use PeerWise for common errors

use PeerWise for smaller examples

practise makes perfect …

# Revision

## know how to study

how do you study for this course?

3. normalisation
4. data mining
5. transactions

use PeerWise for definitions

use PeerWise for examples

use Weka to solve apriori problems

# Questions?

# Good Luck!

on both project and exam

no more lectures, last tutorial today,
group presentations in respective slots on Tuesdays