INFO90002
DATABASE SYSTEMS & INFORMATION MODELLING

Exam Duration: 2 Hours

This paper has 6 pages

Authorised Materials: No materials are authorised.

Instructions to Invigilators:

This exam paper must be handed in at the conclusion of the exam.

Answers are to be written in standard exam script books.

Instructions to Students:

This exam paper has 8 questions, some with multiple parts.

You should answer all questions.

Answers must be written in blue or black ink.

Textual answers can be in point form.

The total mark for this exam is 60 marks, representing 60% of your final assessment. Since you have 2 hours writing time, you should spend about 2 minutes per mark.

Q1 – Database Design (16 Marks)

Scenario:

A newspaper is setting up a website where people can write, read and comment on news stories. Your job is to design a database that can record the information needed for the website to work.

The website lists a number of stories, each classified within one of about 10 sections ('Local news', 'World news', 'Opinion', 'Sport', 'Technology' etc). Each content of each story consists of a piece of text (the manager would prefer there to be no limit on the size) as well as a headline and a short "lede" (a lede is a sentence of 10 to 20 words that summarizes the story). Each story is written by one or more of our authors, who submit the story to our database on a particular date. If a story is considered worthy, it is edited by one of our editors, assigned to a section, and then published on a particular date. Authors and editors are staff members. It is not possible to be both an author and an editor. About each staff member we store their name and when they join (and later leave) the newspaper.

We want to organize the website so that readers can click on a section, or an author's name, and see a list of all the relevant stories. On the main page we list each story's title (shown as a headline) and lede. Then if the reader clicks on the headline, we display the entire story.

Our readers, if they wish, can choose to register themselves in our database, recording a username and password (they do not need to record their real name). Readers who register can then comment on stories, and on other readers' comments. They can also click "Like" on a story – and if they change their mind later, they can "Unlike" it. When we display a story we show the number of Likes the story has received and list out the comments below it. With each comment we show the username of the person who commented, and the time they commented. We aim for brevity in comments, and restrict them to 1024 characters.

Question 1 instructions:

For the above scenario, draw a *logical* data model. (There is no need to determine the data-types of columns.)

Use crows-foot notation for relationships, and join the lines to the related columns. There is no need to add names to relationships.

You do not need to write a data dictionary.

If you wish, you may explain the reasoning behind any design decisions or assumptions you made.

Note that marks are largely based on a workable model that caters to all the constraints stated in the case. Marks may be lost for incorrect entities, relationships, cardinality or notation, lack of detail or internal contradictions.

Q2 – SQL (14 Marks)

Consider the following data model and *sample* data (not all data is shown) for a student real-estate system.

At any given time, an individual property is empty or occupied by one tenancy. Note that the rent charged for a particular tenancy, especially one in the past, may be different to the current advertised rent on that property.
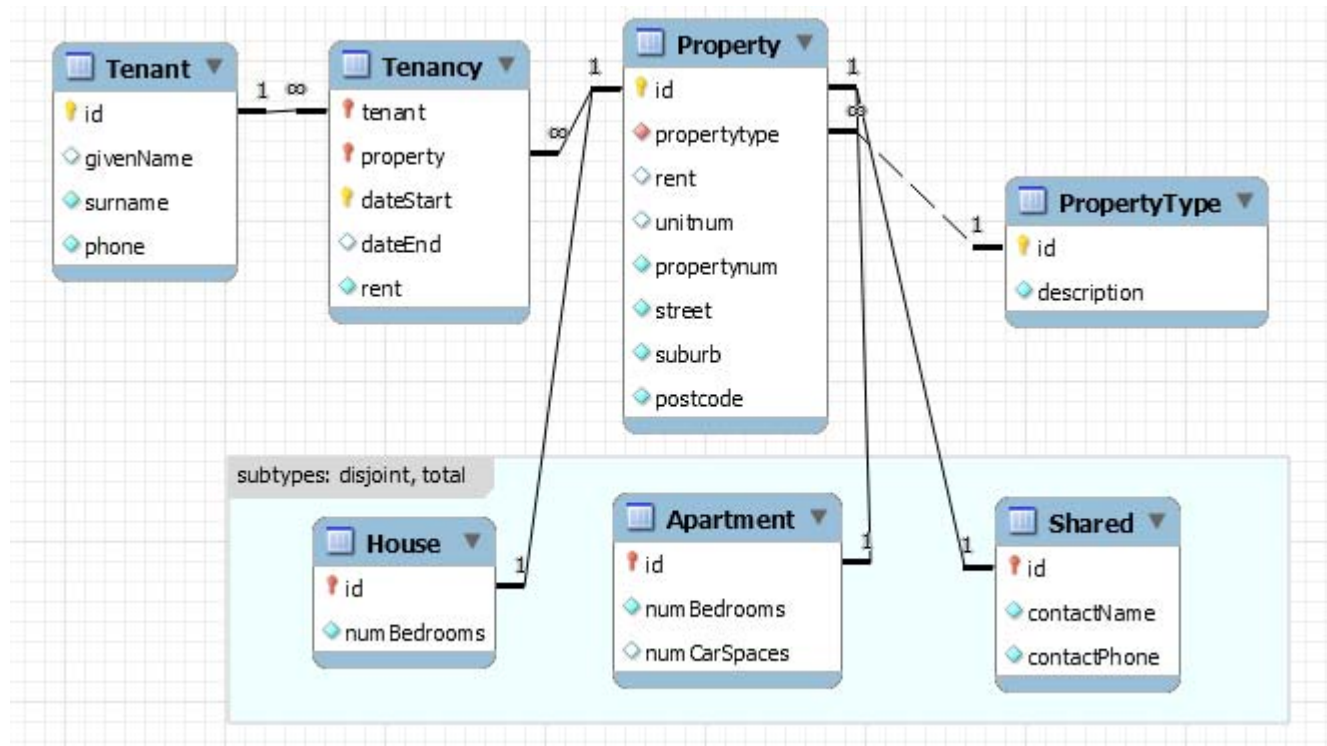
id	givenName	surname	phone
1	Britteny	Abston	0441532370
2	Margy	Alter	0810478134
3	Freida	Amaral	072681647
4	Kera	Anderson	0266223748
5	Barbra	Basham	0353849916
6	Lon	Batchelor	0924374652
7	Heide	Belew	0987544272
8	Susan	Bergen	083194347
9	Ouida	Binion	0599985340
10	Wai	Bisceglia	0606937760

Tenant	Property	dateStart	dateEnd	rent
92	5	2014-08-09	NULL	429
160	13	2014-04-23	2016-04-22	362
153	39	2014-01-11	2015-01-11	339
48	42	2014-05-25	2015-05-25	291
93	48	2014-06-14	NULL	314
79	49	2014-02-18	2015-02-18	517
36	51	2014-01-05	2015-01-05	395
132	51	2014-08-11	NULL	425
6	82	2014-06-06	2015-06-06	384
86	123	2014-04-07	2015-04-07	458
109	133	2014-03-30	2015-03-30	534

id	PropertyType	rent	unitnum	propertynum	street	suburb	postcode
1	H	394	NULL	39	Napier lane	Fitzroy	3065
2	H	404	NULL	97	Charles place	Fitzroy	3065
3	H	540	NULL	68	Harrison place	Fitzroy	3065
4	H	590	NULL	48	Royal lane	Fitzroy	3065
5	H	429	NULL	38	Johnston street	Fitzroy	3065
6	A	374	4	45	Rodney place	Carlton	3053
7	H	586	NULL	12	Bouverie street	Carlton	3053
8	S	107	NULL	26	Kerr street	Fitzroy	3065
9	A	378	4	93	Highett place	Fitzroy	3065
10	S	171	NULL	90	Charles place	Fitzroy	3065

id	description
A	Apartment
H	House
S	Shared

id	numBedrooms
1	2
2	3
3	4
4	4
5	3
7	4
11	3
13	2
14	3
17	3

id	numBedroms	numCarSpaces
6	1	2
9	1	1
12	1	0
15	1	2
21	1	1
24	2	1
30	2	2
33	1	0
36	1	0
39	1	1

id	contactName	contactPhone
8	Finley	8697660699
10	Jovanny	1597228831
16	Deborah	2319846495
20	Dean	8566163983
28	Stanley	9176265253
52	Krish	4854635066
64	Haiden	6938385335
76	Alvaro	7457141435
88	Dylan	6118526965
106	Destiny	7316239255

Tasks:

Questions 2A-D require you to write SQL statements. Your code should be formatted so it is easy to understand. Ensure user friendly output by displaying and sorting on appropriate fields and naming fields appropriately.

A) Which tenants have paid us the most rent?
   List the top 5 payers, and how much rent each has paid in total.   (4 Marks)

B) What is the longest given name among our tenants?
   If several names tie for first place, list them all.   (1 Mark)
   LENGTH(); MAX()

C) List the addresses of the houses and apartments
   which only have one bedroom.   (3 Marks)

D) List the names of tenants who have rented
   at least two different types of properties.   (6 Marks)

Q3 – Normalisation (5 Marks)

We are modelling a database for storing students and their marks. Students have an id number and name, and receive marks for a series of assignments, each of which has in id number and a title. A student does a given assignment only once – we record the date that the student submitted the assignment.

Our modeller has arrived at the following relation (primary key is underlined).

STUDENTMARKS (studentid, givenName, surname (assignmentId, assignmentTitle, mark, dateSubmitted))

But this relation is not in third normal form.

First, explain why it is not.   (1 mark)

Second, convert the relation to 3rd normal form.
Mark your primary keys with a solid underline, and your foreign keys with a dotted underline.

You don't need to show intermediate normal forms – just the 3rd normal form you end up with.   (4 marks)

Q4 – Physical Design (5 Marks)

For each of the following columns in the database shown in question 2,
write the column's MySQL data type, include the width of the column where appropriate

i)   Tenant.id   BIGINT
ii)   Tenant.surname   VARCHAR(15)
iii)   Tenant.phone   CHAR(10)

iv)   Tenancy.tenant   BIGINT
v)   Tenancy.dateEnd   DATE

vi)   PropertyType.id   CHAR(1)

vii)   Property.unitnum   TINYINT
viii)   Property.postcode CHAR(4)
ix)   Property.rent   SMALLINT

x)   Apartment.numBedrooms TINYINT

Q5 – Distributed Databases (7 Marks)

A) You are designing a database for a new social media startup.
Give three (3) reasons why you might choose a distributed database. (3 Mark)
Good fit for distributed organizations; Supports data sharing; Supports DB recovery

B) Describe the differences between replication and partitioning. (2 Marks)
Data copied across sites; Table rows & columns distributed across sites;

C) List four (4) possible disadvantages of a using a distributed database. (2 Marks)
many server sites → higher chance of breach;Lack of standards; Increased training costs: complex IT infrastructure; Increased storage requirements

Q6 – Transactions (8 Marks)

A) Explain, using a timeline diagram and text, how 2 processes simultaneously accessing the same data can create the **Lost Update** problem. (4 Marks)

B) Explain two different purposes for using transactions. (4 Marks)
User need the ability to define a unit of work; Concurrency access to data.
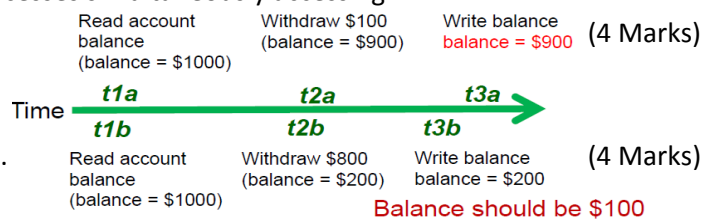
Q7 – Database Administration (3 Marks)

Consider the database in question 2. Our analysis shows that at go-live we will have about 3,000 properties and 5,000 tenants in our database. We expect to register 100 new properties and 500 new tenants per year. About 1,000 new tenancies begin each year.

For each of the tables Tenant, Tenancy, Property and PropertyType, show:

- The number of rows at go-live
- The number of rows after one year
- The number of rows after ten years

Write your answer in a table like this:

	Rows at go-live	Rows at 1 year	Rows at 10 years
Tenant			
Tenancy			
Property			
PropertyType			

Q8 – NoSQL (2 Marks)

Explain why, when implementing a replicated NoSQL database,
we must trade-off between availability and consistency.

According to CAP theorem, we can only pick two out of three of Consistency, Availability, Partition tolerance.

END OF EXAM

Library Course Work Collections

Author/s:

Computing and Information Systems

Title:

Database Systems and Information Modelling, 2016 Semester 1, INFO90002

Date:

2016

Persistent Link:

http://hdl.handle.net/11343/160140