MSc Business Analytics Online Examinations 2019/2020

For internal Students of Imperial College London.
This paper also forms part of the examination for the Associateship.

Fundamentals of Database Technologies
(BUSI97354)

Thursday 2nd July 2020; 13:00-15:00

CLOSED BOOK

Instructions

Answer **ALL** questions and their sub-questions from both Part A and Part B.

Assume you are using the Postgres dialect of SQL.

College approved calculators may be used.

The supplied Postgres information sheet may be used.

Available marks are indicated next to each question. **Total marks available: 50**

This time-limited remote assessment has not been designed to be open book. We expect you not to use any additional resources when completing this assessment. The concept of academic integrity is fundamental to your student experience at Imperial so as to ensure that your academic achievements are a true reflection of your abilities. Academic integrity means conducting all aspects of your academic life in a professional manner. The use of the work of another student, past or present, constitutes plagiarism. Giving your work to another student to use (other than in a group assessment) may also constitute an offence. The College will investigate all instances where an examination or assessment offence is reported or suspected, using plagiarism software, vivas and other tools, and apply appropriate penalties to students who are found guilty. Collusion is a form of plagiarism and will be treated in a similar manner. This is an individual assessment and thus should be solely based on your work.

©Imperial College London 2019/2020

Part A: Theory   (Total: 20 marks)

Question 1

Define a) candidate keys, b) primary keys, c) proper subsets of candidate keys.

(3 marks)

Question 2   (Total: 2 marks)

a) Give, concisely, an example scenario in which the same piece of information is stored in multiple tables and an UPDATE statement causes data corruption. (1 mark)

b) Explain how normalisation prevents this error.   (1 mark)

Question 3   (Total: 2 marks)

a) State the definition of first normal form (1NF).   (1 mark)

b) Does 1NF prevent the storage of an address (including house number, street name, city and postcode) in a single table cell?   (1 mark)

Question 4   (Total: 6 marks)

Consider the following table:

meetings

Attribute name	Description	Datatype
meeting_name	Meeting name	varchar(512)
meeting_id	Meeting ID	real
Meeting_Room	Location	boolean
duration	Meeting length	date
PhoneNumber	Remote working number	string

a) What problems exist concerning the attribute names?   (1 mark)

b) What problems exist concerning the attribute datatypes?   (5 marks)

Question 5

Describe two major problems with data storage before the introduction of relational databases, and explain how relational databases have addressed or fixed these problems.   (4 marks)

Question 6

Briefly define the difference between INNER JOIN and LEFT OUTER JOIN, indicating whether each one is symmetric.   (3 marks)
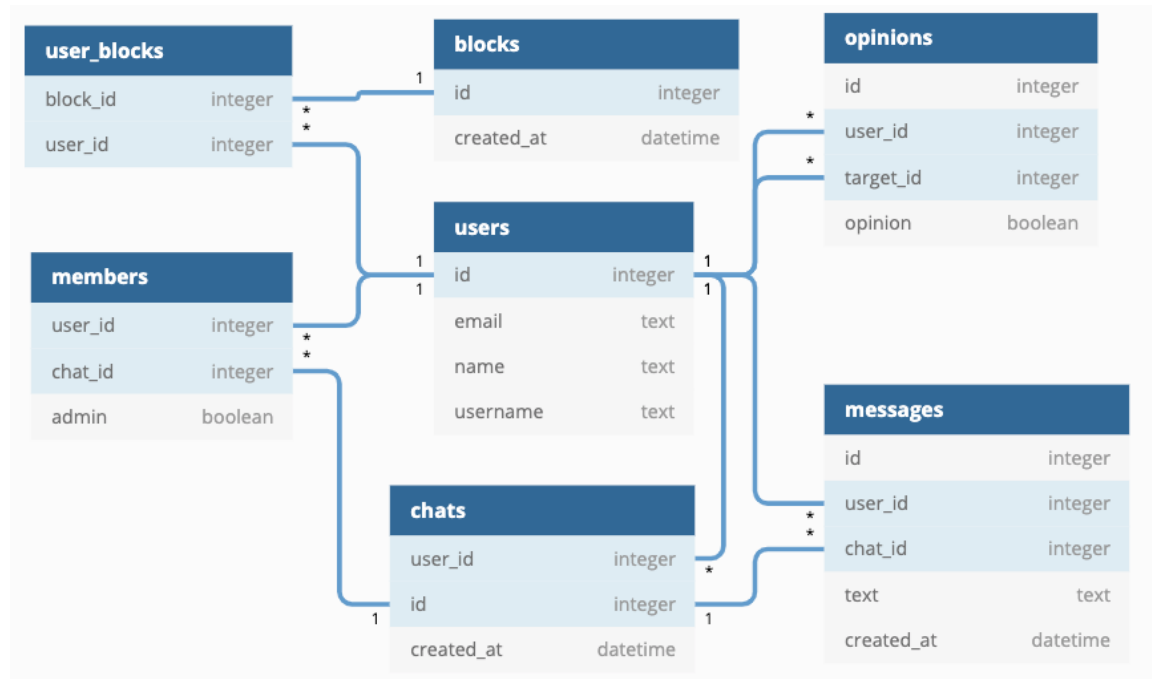
Part B: Queries   (Total 30 marks)

Consider the following entity-relationship (ER) diagram.

Legend
1   Up to one entity at this end of the relationship
*   Unlimited entities at this end of the relationship

This ER diagram depicts a messaging app. The **users** table describes users of the app. The **chats** table describes chats in which multiple users can take part; each chat is also owned by a single user. The **messages** table describes individual messages within chats. The **members** table records which users are allowed to access each chat.

The **blocks** table lists blocks (rules which say that users cannot contact each other). The **user_blocks** table shows which users are part of each block. For example, the following rows would be added to describe two users (IDs 333 and 444) blocking each other:

blocks

id	created_at
459	Tue Jan 17 16:00:00

user_blocks

block_id	user_id
459	333
459	444

The **opinions** table describes users' opinions of each other. A row in the opinions table has user_id, target_id and opinion. If a user has a positive opinion of the target, opinion is TRUE; otherwise, opinion is FALSE.

Question 1

Are there any many-to-many relationships in this ER diagram? Justify your conclusion.

(2 marks)

Question 2

Is the opinions relationship symmetric? In other words, if user A has a positive opinion of user B, is the converse necessarily true? Explain your answer. (2 marks)

Question 3

Write a query to show the full users table, with an additional column showing the number of messages posted by each user. (2 marks)

Question 4

Write a query to show all chats along with the number of messages they contain, ordered chronologically. (3 marks)

Question 5

Write a query to show the time of the most recent message from each user. (2 marks)

Question 6

Write a query to show the name and ID of each user along with the text of their most recently posted message. (3 marks)

Question 7 (Total: 3 marks)

a) Assume that there is a single user who has the most messages. Write a single query to show this user's name, email and number of messages. (2 marks)

b) Explain what will happen if your query is run and there are two users who are tied for the most messages. (1 mark)

Question 8

The LENGTH() function shows the length of a string (its character count).

Write a query to show the messages belonging to user ID 777, with a column showing the character count of each message and a column showing the cumulative character count. Make sensible choices about ordering and make sure that in the displayed result, the cumulative character count increases monotonically. (5 marks)

Question 9

Do users who are involved in blocks receive more negative opinions? Write a query which shows each user ID, along with the number of blocks they are involved in, and the number of negative opinions they have received.   (4 marks)

Question 10   (Total: 4 marks)

Imagine that the full database for the messaging app has been leaked onto the Web.

a) Is it possible to find out who has blocked user ID 777? Explain why. If possible, write a query to do this.   (2 marks)

b) Is it possible to find out who has a negative opinion of user ID 777? Explain why. If possible, write a query to do this.   (2 marks)

End of paper