

Midterm 2

Ques 1. [2 Marks] Circle the correct answers

1. Normalization is not for
 - a) reducing the number of tables in the schema.
 - b) eliminating uncontrolled redundancy of data.
 - c) eliminating anomalies that could otherwise occur with insert, delete or update.
 - d) ensuring that functional dependencies are enforced.

2. If a non-key attribute of a relation is null, then that table automatically violates which normal form
 - a) NONE
 - b) 1NF
 - c) 2NF
 - d) 3NF
 - e) BCNF

3. A table is in 2NF if the table is in 1NF and what other condition is met?
 - a) There are no functional dependencies.
 - b) There are no null values in primary key fields.
 - c) There are no repeating tuples.
 - d) There are no attributes that are not fully functionally dependent on the relation's primary key.

4. Which is true?
 - a) A superkey is a key.
 - b) Two tuples can have the same key value.
 - c) A relation can have more than one candidate key.
 - d) None of the above

Normalization, 3NF, BCNF

Ques 2. [4 Marks] Consider the relation shown

A	B	C
A1	B1	C1
A1	B1	C2
A2	B1	C1
A2	B1	C3

a) List all the functional dependencies that the given relation instance satisfies.

b) Assume that the value of attribute C of the last record in the relation is changed from C3 to C2. Now list all the functional dependencies that this relation instance satisfies.

Ques 3. [2 Marks] Consider a relation $R = (A, B, C, D, E)$.

R is decomposed it into $R_1 = (A, B, C)$, $R_2 = (A, D, E)$. The set of functional dependencies is:

$A \rightarrow BC$

$CD \rightarrow E$

$B \rightarrow D$

$E \rightarrow A$.

Show that this decomposition is a lossless-join decomposition.

Ques 4. **[6 Marks]** Consider the below instance of the CLASS relation. Determine whether each of the following statements is true or false; briefly justify your answer.

<u>CID</u>	CName	Credit	<u>SID</u>	SName	Grade
C1	Math	4	S1	Jimmy	Pass
C1	Math	4	S2	Sean	Pass
C2	AI	4	S3	Mary	Pass
C3	DB	3	S4	Peter	Fail
C3	DB	3	S5	Lila	Pass
C3	DB	3	S6	Tony	Pass
C4	SE	3	S1	Jimmy	Pass
C4	SE	3	S2	Sean	Fail
C5	DM	3	S7	Boi	Pass

Statements	T/F?	Brief Justification
a) If class (C5, DM) is deleted then the information about student S7 would also be removed because of deletion anomaly.		
b) Inserting CLASS (C6, Calculus, 3) is an example of insertion anomaly.		

Ques 5. **[6 Marks]** Consider the following two schemas

Schema 1: R (A,B,C,D)

Schema 2: R1 (A,B,C) and R2 (B,D)

Determine whether each of the following statement is true or false; briefly justify your answer.

Statements	T/F?	Brief Justification
a) Consider Schema 1 and suppose that the only functional dependencies that hold on the relations in this schema are $A \rightarrow B$ and $C \rightarrow D$. Is Schema 1 in Boyce-Codd Normal Form (BCNF)?		
b) Consider Schema 2 and suppose that the only functional dependencies that hold on the relations in this schema are $A \rightarrow B$, $A \rightarrow C$, $B \rightarrow A$, $A \rightarrow D$ and all possible implicit FDs. Is Schema 2 in BCNF?		
c) If $A \rightarrow D$ is deleted from part (b). Is Schema 2 in BCNF?		

Ques 6. **[15 Marks]** Consider the following relation $R(A, B, C, D)$ and the functional dependencies that hold over this relation.

$F = A \rightarrow B, D$

$C, D \rightarrow B$

$C \rightarrow D$

$B \rightarrow D$

a) Determine all candidate keys of R .

b) Compute the closure of $\{C, B\}$

c) Compute the minimal cover

- d) Is R in 3NF? If not, then decompose it. Show all the steps. If you are decomposing R then determine the candidate keys of each decomposed relations.

SQL

Ques 7. **[6 Marks]** Consider the below relations where primary keys are underline and foreign key is bold.

employee (ssn, name, address, phone)

project (pno, pname, **essn**)

Determine whether each of the following statement is true or false; provide a brief justification of your choice. State any assumptions that you make.

Statements	T/F?	Brief Justification
a) In SQL, the query "SELECT name FROM employee where ssn>10" will not return duplicates.		
b) Consider below queries: Query 1: SELECT COUNT (*) FROM employee Query 2: SELECT COUNT(name) FROM employee Query1 and Query2 always return the same result.		
c) SELECT e.name FROM employee e, project p WHERE e.ssn = p.essn is equivalent to SELECT name FROM employee		

Ques 8. **[5 Marks]** Give the SQL command needed to create the Item relation. You may choose the reasonable domains for your attributes. You must specify the following:

- ItemNo is the primary key
- {IName, Quantity} is a candidate key
- IName and Quantity are not allowed to contain null values
- Price must be greater than 100.

Ques 9. **[6 Marks]** Consider the following relations (Keys are underlined and foreign keys are bold)

Author (Name, Address, Gender, Research_area)

Journal (JName, Paper_title, Published_year, **AName**)

Write SQL statements that correspond to the following relational algebra expressions.

$$\pi_{JName, Paper_title}(\sigma_{Published_year=2016}(Journal))$$
$$\pi_{JName}(\sigma_{Aname=name \text{ AND } published_year=2016}(\text{Author X Journal}))$$

Ques 10. **[24 Marks]** Consider the Company database schema and instances given in Appendix A. Use SQL to answer the following questions. Please make sure that your queries do not include any extra information.

a) For each project, retrieve the project number, the project name, and the number of employees from department 5 who work on the project.

b) Retrieve the name of all employees who have two or more dependents. **(Use nested query)**

c) Create a view that has employee name, supervisor name, and employee salary for each employee who works in the 'Research' department

d) Retrieve all employees whose address is in Houston, Texas

e) Retrieve the list of all project numbers for projects that involve an employee whose name is 'John Smith', either as a worker or as a manager of the department that controls the project.

f) Retrieve the name of all employees who works on all projects controlled by department number 5.

Datalog

Ques 11. **[12 Marks]** Consider the following relations (Primary keys are underline and foreign keys are bold):

Paper(title, citation, **AName**, year, **JName**)

Journal(JName, year, PName, **AName**)

Author (AName, RArea)

Where Aname is author name, PName is publisher name and RArea is research area.

Write Datalog to answer the following queries.

a) List the journal name and author name of all journals in the database.

b) List the name of journals in which papers are published by authors having research area "Databases"

A continuation of the above question; the relations are listed here again to prevent having to look them up

Paper(title, citation, **AName**,year,**JName**)

Journal(JName, year, PName, **AName**)

Author (AName, RArea)

Where Aname is author name, PName is publisher name and RArea is research area.

c) List the citation of all paper authored in 2016 or authored by Hazra Imran

d) Find the names of authors who have authored paper for all journals.

This space is intentionally left blank. You can use it to answer questions or as scratch paper

APPENDIX A

Name: _____ Student ID _____

Consider the relation schema and instance of the COMPANY database given below:

