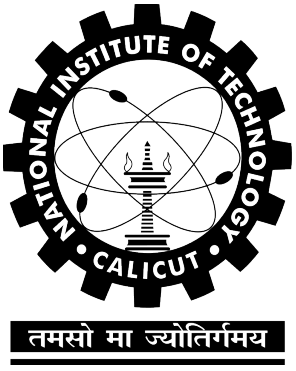
CS3002 : Database Management System

A
Course File
By

Nadiya T T



Department of Computer Science and Engineering

National Institute of Technology, Calicut

Winter-2017

Table of Contents

1 Course Plan 3

2 First Mid-term Question Paper 5

3 First Mid-term Marks Sheet 7

4 Second Mid-term Question Paper 9

5 Second Mid-term Marks Sheet 11

6 End-Semester Question Paper 13

7 Course Outcome Attainment Scores 17

Department of Computer Science and Engineering

National Institute of Technology Calicut

Tentative Course Details – Monsoon Semester 2016

CS3002 DATABASE MANAGEMENT SYSTEMS

Pre-requisite: Nil

Course Outcomes:

CO1: Model, design and normalize databases for real life applications.

CO2: Code and deploy databases for applications using RDBMS like Mysql

CO3: Query Database applications using Query Languages like SQL

CO4: Undertake and successfully complete Database Development projects within the allotted time.

CO5: Deploy efficient IT solutions using free and open software and help the society

About the Lecture:

Lecture Hours (A/A+ slot): Mon (8-9AM), Tue (1-2PM),Wed (9-10AM)and Fri (10.15-11.15AM)
Lecture Hall: MB 207

About the Instructor:

Name: Nadiya T T
Room: MB 209E
Email: nadiyasabin@nitc.ac.in

References:

1. Ramez Elmasri and Shamkant B. Navathe, Fundamentals of Database Systems (5/e), Pearson Education, 2008.
2. Raghu Ramakrishnan and Johannes Gehrke, Database Management Systems (3/e), McGraw Hill, 2003.
3. Peter Rob and Carlos Coronel, Database Systesm- Design, Implementation and Management (7/e), Cengage Learning, 2007.

Split up of marks:

MidTerm I	:	15 marks
MidTerm II	:	15 marks
Assignment/Quiz	:	20 marks
End Exam	:	50 marks

Course Schedule:

Sl. No.	WEEK #	TOPIC
1	WEEK 1	Database System concepts and architecture, Data modeling using Entity Relationship (ER) model
2	WEEK 2	Specialization, Generalization, The Relational Model
3	WEEK 3	Relational database design using ER to relational mapping, Domain Relational Calculus
4	WEEK 4	SQL, Relational algebra
5	WEEK 5	Relational calculus, Tuple Relational Calculus
6	WEEK 6	Test 1
7	WEEK 7	Database design theory and methodology, Functional dependencies
8	WEEK 8	Normalization of relations, Normal Forms
9	WEEK 9	Data Storage and indexing, Single level and multi level indexing
10	WEEK 10	Dynamic Multi level indexing using B Trees and B+ Trees
11	WEEK 11	Test 2
12	WEEK 12	Transaction processing concepts, Schedules
13	WEEK 13	Serializability, Concurrency control
14	WEEK 14	Two Phase Locking Techniques, Optimistic Concurrency Control
15	WEEK 15	Database recovery concepts and techniques, Introduction to database security
16	WEEK 16	Properties of relational decomposition, Algorithms for relational database schema design

Roll No: ..... Name .....

National Institute of Technology Calicut

Department of Computer Science & Engineering

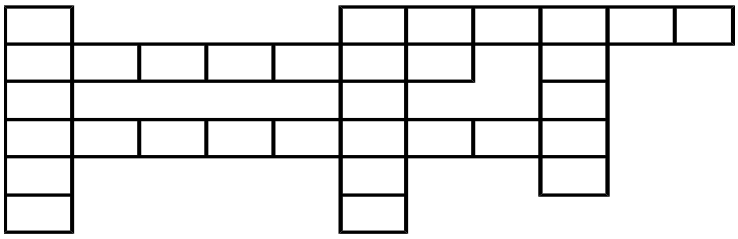
CS3002 Database Management Systems

First MID Term Exam (Monsoon Semester 2016)

Max. Marks: 15 Time: 1hr

---

1. Solve the following [1.5]



Down: 1. The overall design of the database.

2. The set of all allowable values that attribute may assume.

3. A subset of a database that is generated from a query

Right: 2. Number of attributes in a relation

4. A directory of information about data sets, files, or a database

5. The data stored in database at a particular moment of time

2. Consider the following relations for a database that keeps track of business trips of salespersons in a salesoffice:

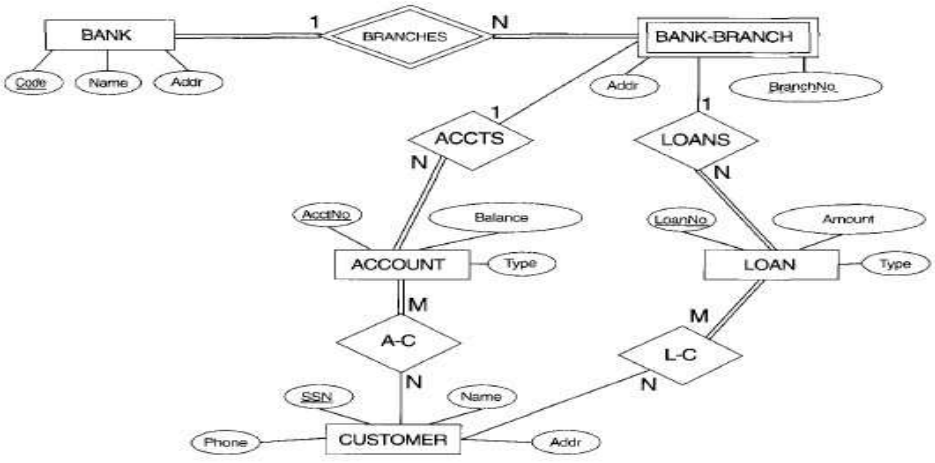
SALESPERSON (SSN, Name, Start_Year, Dept_No)

TRIP (SSN, From_City, To_City, Departure_Date, Return_Date, Trip_ID)

EXPENSE (Trip_ID, Account#, Amount)

Specify the foreign keys for this schema, stating any assumptions you make. [1]

3. Consider the ER diagram shown in Figure for part of a BANK database.



Each bank can have multiple branches, and each branch can have multiple accounts and loans.

a. List the (nonweak) entity types in the ER diagram. [.25]

b. Is there a weak entity type? If so, give its name, partial key, and identifying relationship. [.25]

c. What constraints do the partial key and the identifying relationship of the weak entity type specify in this diagram? [1]

d. List the names of all relationship types, and specify the (min, max) constraint on each participation of an entity type in a relationship type. Justify your choices. [1]

e. List concisely the user requirements that led to this ER schema design. [1]

4. Define DDL,DML,SDL and VDL [2]

5. Consider the following two aspects and Modify the BANK schema using ER and EER concepts of specialization and generalization. [2]

- There are three different kinds of ACCOUNTs namely SAVINGS_ACCTS, CHECKING_ACCTS, and TRUSTS. For each ACCOUNT we have to take care of its TRANSACTIONs. Each TRANSACTION has a *type*(such as 'deposit', 'withdrawal' or 'check'). Furthermore, each TRANSACTION has a *date/time*(consisting of date and time) and an *amount*.
- There are different kinds of LOANS, namely CAR_LOANS, HOME_LOANS, CREDIT_LINE and PERSONAL ones. For each LOAN we have to take care of its PAYMENTs. Each PAYMENT has a *type*, *date* and *amount*.

6. Consider the following schema: [2]

Suppliers(sid: integer, sname: string, address: string)

Parts(pid: integer, pname: string, color: string)

Catalog(sid: integer, pid: integer, cost: real)

State SQL queries correspond to the following queries

1. $(\pi_{sname}((\sigma_{color=red}\text{Parts})\bowtie(\sigma_{cost<100}\text{Catalog})\bowtie\text{Suppliers}))\cap$
$(\pi_{sname}((\sigma_{color=green}\text{Parts})\bowtie(\sigma_{cost<100}\text{Catalog})\bowtie\text{Suppliers}))$
2. $\pi_{sname}((\pi_{sid,sname}((\sigma_{color=red}\text{Parts})\bowtie(\sigma_{cost<100}\text{Catalog})\bowtie\text{Suppliers}))\cap$
$(\pi_{sid,sname}((\sigma_{color=green}\text{P arts})\bowtie(\sigma_{cost<100}\text{Catalog})\bowtie\text{Suppliers})))$

7. Consider the database. [3]

Flights(flno: integer, from: string, to: string,distance: integer, departs: time, arrives: time)

Aircraft(aid: integer, aname: string, cruisingrange: integer)

Certified(eid: integer, aid: integer)

Employees(eid: integer, ename: string, salary: integer)

Give an expression in the relational algebra for each of the following :

a. Find the aids of all aircraft that can be used on non-stop flights from Bonn to Madras.

b. Find the names of pilots who can operate some plane with a range greater than 3,000 miles but are not certified on any Boeing aircraft

c. Find the eids of employees who make the highest salary.

National Institute of Technology Calicut
Statement of Test Result I for Monsoon Semester 2016-2017

Stream : MASTER OF COMPUTER APPLICATIONS
Semester : III
Class Batch : -ALL-
Subject : CS3002 - DATABASE MANAGEMENT SYSTEMS
Total Classes : 37
Faculty : CS9 - (ADHOC)
Total Marks : 15

Sr. No	Reg. No.	Student Name	Marks	Class Attended	% Marks	% Atten.
1	M150038CA	ABHISHEK SUMEET TOPPO	6	0	40	0
2	M150059CA	AJAY YADAV	9	0	60	0
3	M150495CA	AKSHITA PANDEY	7.75	0	51.67	0
4	M150483CA	ALKA KUMARI	10	0	66.67	0
5	M150043CA	AMIT KUMAR	10.75	0	71.67	0
6	M150479CA	AMIT KUMAR SONI	6	0	40	0
7	M150057CA	ANIKET SHARMA	8	0	53.33	0
8	M150486CA	APURVA KUMAR SINGH	12.5	0	83.33	0
9	M150503CA	AVNISH AGRAHARI	8	0	53.33	0
10	M150481CA	BISHAKHA KUMARI	10.25	0	68.33	0
11	M150045CA	DIWAKAR PRAJAPATI	1.5	0	10	0
12	M150061CA	HARSH VERMA	10.75	0	71.67	0
13	M150062CA	JIWAN PRAKASH CHOUDHARY	5	0	33.33	0
14	M150039CA	JYOTI PANDEY	8.5	0	56.67	0
15	M150037CA	LAKSHIT GARG	9.25	0	61.67	0
16	M150538CA	LAVAKUSH MANI TIWARI	8	0	53.33	0
17	M150060CA	MANISHA JAISWAL	13	0	86.67	0
18	M150053CA	MAYANK VELANKAR	11	0	73.33	0
19	M150052CA	MOHAMMAD ARSH	7	0	46.67	0
20	M150511CA	MUSHAHID KHAN	4.5	0	30	0
21	M150041CA	NAMITA KESHARWANI	5.5	0	36.67	0
22	M150048CA	PANKAJ KUMAR	1.25	0	8.33	0
23	M150058CA	PRATYUSH AGARWAL	6.75	0	45	0
24	M150034CA	PRIYANKA EKKA	7.5	0	50	0
25	M150050CA	RAHUL KUMAR	3.5	0	23.33	0

26	M150049CA	RAUSHAN KUMAR	2.75	0	18.33	0
27	M150512CA	RAVINDRA CHANDRA JADALI	5.5	0	36.67	0
28	M150054CA	SARVESH SINGH	3.75	0	25	0
29	M150040CA	SAURABH CHHABRA	4.75	0	31.67	0
30	M150487CA	SHAH FAHAD	6.25	0	41.67	0
31	M150036CA	SHIVAM AGARWAL	5.25	0	35	0
32	M150047CA	SOMNATH SAMANTA	2.5	0	16.67	0
33	M150035CA	SOORYA E	6	0	40	0
34	M150042CA	SUNNY	8.75	0	58.33	0
35	M150419CA	VAIBHAV PANT	8.25	0	55	0
36	M150480CA	YOGENDRA	8.75	0	58.33	0

Roll No: ..... Name .....

National Institute of Technology Calicut

Department of Computer Science & Engineering

CS3002 Database Management Systems

Second MID Term Exam (Monsoon Semester 2016)

Max. Marks: 15 Time: 1hr

---

1. State whether the following conclusion are true or false: [1]

   a. NOT (P(x) OR Q(x)) $\Rightarrow$ (NOT (P(x)) AND (NOT (Q(x)))

   b. NOT ($\exists$ x) (P(x)) $\Rightarrow$ $\forall$ x (NOT (P(x))

2. Identify functional dependencies in the following table [1]

A	B	C
a1	b1	c1
a1	b1	c3
a1	b2	c1

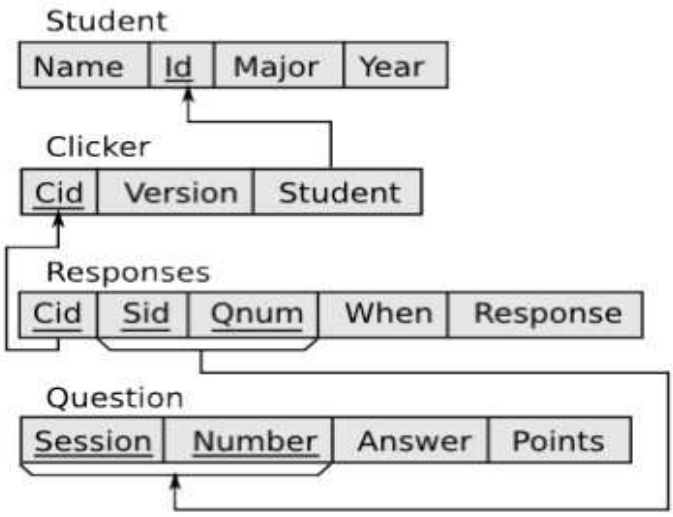
3.Show how you can specify the following relational algebra operations in both tuple and domain relational calculus. [3]

   a. R(A, B, C) MINUS S(A, B, C)

   b. R(A, B) DIVIDE S(A)

4. Consider the Clicker database shown in figure. Express the given query in SQL, relational algebra, Tuple relational calculus and Domain relational calculus [3]

"List the IDs and version numbers of those clickers used to give at least one correct answer during the '24-A-2' session but that are not registered to any student."



5. State and Prove pseudotransitive rule of Functional Dependencies. [2]

6. Consider the relation R= { Course_no, Offering_deptNo, Offering_deptName, Instructor_ssn, Semester, Year,Address, Room_no, Building_Name, Room_Size }

Besides you have the following functional dependencies:

{ Course_no} -->{ Offering_deptNo}

{ Offering_deptNo}-->{ Offering_deptName}

{ Course_no, Semester , Year}--> {Address, Room_no , Instructor_ssn}

{Building_Name} --> {Address}

Try to determine which sets of attributes form keys of R. How would you normalize this relation? [3]

7. Consider the relation : BOOK (Book_Name , Author , Edition , Year) with the data: [2]

Book_Name	Author	Edition	Copyright_Year
DB_fundamentals	Navathe	4	2004
DB_fundamentals	Elmasri	4	2004
DB_fundamentals	Elmasri	5	2007
DB_fundamentals	Navathe	5	2007

a. Is there exist any MVD in the BOOK relation? then list it.

b. If any MVD exist then what would be the decomposition of this relation based on that MVD? Evaluate each resulting relation for the highest normal form it possesses.

National Institute of Technology Calicut
Statement of Test Result II for Monsoon Semester 2016-2017

Stream : MASTER OF COMPUTER APPLICATIONS
Semester : III
Class Batch : -ALL-
Subject : CS3002 - DATABASE MANAGEMENT SYSTEMS
Total Classes : 36
Faculty : CS9 - (ADHOC)
Total Marks : 15

Sr. No	Reg. No.	Student Name	Marks	Class Attended	% Marks	% Atten.	Assignment
1	M150038CA	ABHISHEK SUMEET TOPPO	2.5	0	16.67	0	3.25
2	M150059CA	AJAY YADAV	7.75	0	51.67	0	4.25
3	M150495CA	AKSHITA PANDEY	11.25	0	75	0	3.75
4	M150483CA	ALKA KUMARI	10.25	0	68.33	0	4.00
5	M150043CA	AMIT KUMAR	8	0	53.33	0	4.00
6	M150479CA	AMIT KUMAR SONI	6.75	0	45	0	3.25
7	M150057CA	ANIKET SHARMA	11	0	73.33	0	4.00
8	M150486CA	APURVA KUMAR SINGH	7.75	0	51.67	0	4.25
9	M150503CA	AVNISH AGRAHARI	5	0	33.33	0	3.50
10	M150481CA	BISHAKHA KUMARI	8.25	0	55	0	4.00
11	M150045CA	DIWAKAR PRAJAPATI	3.25	0	21.67	0	4.00
12	M150061CA	HARSH VERMA	7.5	0	50	0	4.50
13	M150062CA	JIWAN PRAKASH CHOUDHARY	5.75	0	38.33	0	3.50
14	M150039CA	JYOTI PANDEY	5.75	0	38.33	0	4.25
15	M150037CA	LAKSHIT GARG	9.75	0	65	0	4.50
16	M150538CA	LAVAKUSH MANI TIWARI	5	0	33.33	0	4.13
17	M150060CA	MANISHA JAISWAL	11.75	0	78.33	0	4.00
18	M150053CA	MAYANK VELANKAR	6.75	0	45	0	4.00
19	M150052CA	MOHAMMAD ARSH	8.5	0	56.67	0	4.00
20	M150511CA	MUSHAHID KHAN	5	0	33.33	0	5.00
21	M150041CA	NAMITA KESHARWANI	5	0	33.33	0	3.75
22	M150048CA	PANKAJ KUMAR	7.25	0	48.33	0	3.50
23	M150058CA	PRATYUSH AGARWAL	6.5	0	43.33	0	4.50
24	M150034CA	PRIYANKA EKKA	7.75	0	51.67	0	4.75
25	M150050CA	RAHUL KUMAR	2.5	0	16.67	0	3.00

26	M150049CA	RAUSHAN KUMAR	5.75	0	38.33	0	2.75
27	M150512CA	RAVINDRA CHANDRA JADALI	4.75	0	31.67	0	2.00
28	M150054CA	SARVESH SINGH	6.75	0	45	0	3.00
29	M150040CA	SAURABH CHHABRA	4	0	26.67	0	4.00
30	M150487CA	SHAH FAHAD	8.25	0	55	0	4.25
31	M150036CA	SHIVAM AGARWAL	8	0	53.33	0	2.50
32	M150047CA	SOMNATH SAMANTA	7	0	46.67	0	2.50
33	M150035CA	SOORYA E	9.5	0	63.33	0	4.75
34	M150042CA	SUNNY	9	0	60	0	4.00
35	M150419CA	VAIBHAV PANT	7.75	0	51.67	0	3.00
36	M150480CA	YOGENDRA	10	0	66.67	0	5.00

Roll No: ..... Name .....

National Institute of Technology Calicut

Department of Computer Science & Engineering

CS3002 Database Management Systems

End Semester Exam (Monsoon Semester 2016)

Max. Marks: 50 Time: 3hr

---

1. Match the following [2.5]

a. Immediate Update	1. READ WRITE
b. Lost update problem	2. NO UNDO/REDO
c. Deferred Update	3. WRITE WRITE
d. Unrepeatable Read problem	4. UNDO/ REDO
e. Dirty Read Problem	5. WRITE READ

2. Consider a database that is read-only. Suppose serializability needs to be supported. Please list(all) correct statements from the given [1]

   a. No locking is necessary.

   b. Only read locks are necessary and they need to be held until the end of transaction.

   c. Only read locks are necessary but they can be released as soon as the read is complete.

   d. Both read and write locks are necessary and locking must be done in two phases.

   e. None of the above.

3. Define the following terms in one sentence: [6]

   a) indexing field b) primary key field c) clustering field

   d) secondary key field e) dense index f) nondense (sparse) index.

4. List four DML commands. [1]
5. Draw the three schema architecture of the database system. [1]
6. What is the star property of Bell La Padula Model? [1]
7. State and prove pseudo transitive rule of FDs [1.5]
8. Given the table below, where A, B, C, D and E represent the attributes of the table. Find the functional dependencies F that hold over this relation. [2]

A	B	C	D	E
a1	b1	c1	d1	e1
a1	b2	c2	d1	e2
a1	b3	c1	d1	e3
a2	b1	c1	d2	e4
a2	b2	c2	d2	e5

9. Draw the E/R diagram for the following databases with participation and cardinality constraints

   a) Design a database for a bank, including information about customer and their accounts.

Information about a customer includes their name, address, phone, and Social Security number. Accounts have numbers, types (e.g, saving, current) and balances. Also record the customer(s) who own an account. [2]

b) Change your diagram so an account can have only one customer. [1]

c) Change your original diagram so that a customer can have a set of addresses (with street-city-state triples) and a set of phone numbers (with area code and number). [1]

10. How do optimistic concurrency control techniques differ from the other concurrency control techniques? Why are they also called validation or certification techniques? [2]

11. Consider the two tables T1 and T2. Show the results of the following operations: [3]

Table T1

A	**B**	**C**
d	s	3
d	t	1
e	s	5
f	u	7

Table T2

J	**K**	**L**
h	t	9
f	s	3
g	x	2
d	v	8

a) Select A, J, K from T1 , T2 where A=J or B=K;

b) $T1 \bowtie_{A=J \text{ AND } C<L} T2$

c) {a.B, a.C, b.L | T1(a) ˆ T2(b) ˆ a.A = ‘e’ˆ a.B ≠ b.K}

12. For each of the following schedules determine which properties this schedule has. E.g., a schedule may be recoverable and cascade-less. Draw and fill the table in your answer sheet, and write the justification (Without proper justification answers won’t consider.)

S1 = r1(C), w1(C), r1(A), w1(A), r2(B), r2(A), w2(B), c2 , w1(C), c1 [6]

S2 = w1(A), r1(B), r3(B), w2(A), r2(B), w1(C), c1 , w3(B), c2 , c3

S3 = r1(A), w1(A), r2(A), w2(A), r3(A), w3(A), r2(B), w2(B), c2 , r1(B), w1(B), c1 , c3

Schedules	**recoverable**	**cascade-less**	**conflict-serializable**	**2PL**
S1				
S2				
S3				

13. Consider the following relation R(A, B, C, D, E) and functional dependencies F that hold over this relation.

F = A, B → C

C → B

A, C → D, E

A, B, D → E

a) Determine all candidate keys of R. [1.5]

b) In which normal form is relation R (recall that a relation can be in multiple normal forms). [1.5]

14. Create a B + Tree of order 3 for the following elements, where elements are inserted in the order

    23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 18, 21, 10, 74, 78, 15, 16, 20, 24, 28

    a) Show how the tree will expand & what the final tree will look like. [2]

    b) Delete 59, 56 & 60 in order. [1]

15. Given the relation schema R = (A, B, C, D, E) and the canonical cover of its set of functional dependencies F = {A → BC

    CD → E

    B → D

    E → A }

    Compute a lossless join decomposition in Boyce-Codd Normal Form for R. Show your steps clearly. [3]

16. A file has r = 20000 *STUDENT* records of fixed length. Each record has the following fields: *NAME* (30 bytes), *SSN* (9 bytes), *ADDRESS* (40 bytes), *PHONE* (10 bytes), *BIRTHDATE* (8 bytes), *SEX* (1 byte), *MAJORDEPTCODE* (4 bytes), *MINORDEPTCODE* (4 bytes), *CLASSCODE* (4bytes, integer), and *DEGREEPROGRAM* (3 bytes). An additional byte is used as a deletion marker. The file is stored on the disk and block size B = 512 bytes; [2]

    a) Calculate the record size R in bytes.

    b) Calculate the blocking factor ***bfr*** and the number of file blocks ***b***, assuming an unspanned organization.

17. Consider the following two transactions: [2]

    T1:
    begin
    write C
    read B
    write C
    commit
    T2:
    begin
    write B
    read C
    read C
    commit

    In a DBMS using the two-phase locking algorithm, the transactions will cause deadlocks depends on how they are executed. If the above two transactions are executed concurrently, under what situations can a deadlock occur?

18. Consider the following database schema and example instance for a flight information system: [6]

property

pId	price	owner	sqrFeet	managedBy	location
1	100,000	Alice	560	Property Pete	Lake View
2	3,400,000	Bob	2,000	Hyde Park Prop	Hyde Park
3	1,200,000	Bob	1,200	Property Pete	Hyde Park
4	5,000,000	Martha	800	Fancy Rentals	Evanston

management

mgmName	Location	yearlyProfit
Property Pete	Lincoln Park	34,000,000
Hyde Park Prop	Downtown	3,000,000
Fancy Rental	Lake View	25,000,000

maintenance

cmpName	empName	salary	Location
SuperPlumbing	George	10,000	Lake View
SuperPlumbing	Dave	30,000	Lake View
Carpeting	Keith	15,000	South Chicago

repairs

cmpName	empName	pId	date	type
SuperPlumbing	George	1	2013-12-12	sink
SuperPlumbing	George	1	2013-12-13	toilet
Carpeting	Keith	4	2012-01-01	paining

a) Write a relational algebra expression that returns the names of maintenance personal (empName) that did repair a property in '***Hyde Park'.***

b) Write a relational algebra expression that returns the number of repairs for property managed by property management company ***'Property Pete'***(cmpName) per location. E.g.,this should return (2, HydePark) if there is a property managed by '***Property Pete'*** in ***Hyde Park*** that has been repaired twice.

c) Write an SQL query that returns all owners whose sink's have been repaired, i.e., that own a property where the sink has been repaired.

d) Write an SQL query that returns average price of properties in ***Lake View*** that are between 500 and 800 square feet large.

e) Write a TRC & DRC expression that returns the pId, owner, and location of all properties that are larger than 600 square feet (sqrFeet).

Course Outcome Attainment Scores

CO1: Model, design and normalize databases for real life applications. : 1.08
CO2: Code and deploy databases for applications using RDBMS like Mysql : 1.3
CO3: Query Database applications using Query Languages like SQL : 2.68
CO4: Undertake and successfully complete Database Development projects : 3

Weighted Average CO Attainment : 1.94
Cumulative Percentage Attainment of COs : 64.61

PO1 : 2.09
PO2 : 2.32
PO3 : 2.32
PO4 : 2.13
PO5 : 2.25
PO6 : 0
PO7 : 0
PO8 : 0
PO9 : 0
PO10 : 0
PO11 : 2.25
PO12 : 2.04

Weighted Average PO Attainment : 1.28
Cumulative Percentage Attainment of POs : 42.79