

# BUS 464-D100 MIDTERM EXAM

#### Date: Oct 13, 2015 Name:

**Instructor: Nilesh Saraf Student Number:**

Instructions (read these now):

* Write your name and student number above in the spaces provided.
* Write your student number at the bottom of all pages in this exam.
* At your desk you may have: pens, pencils, eraser, correcting fluid, bottle of water. No cell phones or any electronic devices are allowed. *This exam is closed books, closed notes.*
* **This exam ends at 4:40 pm**.
* The marks allotted to each question are shown. You should budget your time accordingly.
* Write each answer legibly in the answer booklet. You are allowed to scribble on the exam itself using a pen or pencil. For every answer make it very clear to me where to find your answer.
* If you do not understand a question, make reasonable assumptions, and write them down and work from there. Invigilators may not help you interpret questions.
* You are allowed to separate the papers in the exam but the entire exam needs to be returned. Tuck the exam inside your answer booklet.

#### Caution: In accordance with the Academic Honesty Policy (T10.02), academic dishonesty in any form will not be tolerated. Prohibited acts include, but are not limited to, the following:

* **making use of any books, papers, electronic devices or memoranda, other than those authorized by the examiners;**
* **speaking or communicating with other students who are writing examinations;**

**copying from the work of other candidates or purposely exposing written papers to the view of other candidates.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Section** | **Max Marks** | **Marks** |
|  | 1a | 5 |  |
| 1b | 2 |  |
| **TOTAL** | 1 | 7 |  |
|  | 2a | 10 |  |
| 2b | 5 |  |
|  | 2c | 3 |  |
| **TOTAL** | **2** | 18 |  |
|  | 3a | 5 |  |
| 3b | 5 |  |
| 3c | 5 |  |
| **TOTAL** | **3** | 15 |  |
| **EXAM TOTAL** | | **40** |  |

* In your SQL queries feel free to abbreviate the following statements as follows:

SELECT - SE

FROM - FR

WHERE - WH

* Feel free to use abbreviations for fields too – empname or ename instead of EmployeeName.
* You can create a subquery separately and call it, say Query1. Then you can simply write the outer query as:

Select \*

from Table1, (Query1) as X

WHERE <JOIN Table1 & Query1>

**Question 1 (5 pts)**

1. **Multiple Choice Questions (5 points):**
2. The main point(s) in the nytimes article “For big data scientists janitor work is key hurdle to insights” is (check any one):
   1. The growing demand for computer scientists
   2. Minimize data analytic efforts
   3. The impending redundancy of computer scientists
   4. The multiplicity of data sources
   5. That janitor work will soon be eliminated because of sophisticated tools.
3. In a table, an update anomaly can occur when
   1. inserting data
   2. deleting data
   3. inserting or deleting data
   4. All of the above
   5. None of the above
4. Any M:M relationship can be thought of as
   1. One 1:M relationship and one 0:M relationship
   2. Unary relationship
   3. Binary relationship
   4. One 1:M and one ternary relationship
   5. Two 1:M relationships
5. Unlike 1:M relationship a \_\_\_\_ relationship does not require a foreign key.
   1. M:1
   2. M:M
   3. 1:M
   4. 1:1
   5. None of the above
6. MySQL Workbench is a \_\_\_\_.
   1. Computer
   2. Client
   3. Cloud
   4. Query processor
   5. Server

**Question 1**

1. **Short answer (2 points)**
2. **Consider the following database:**

Customer(Custid, custName) # has 200 customers only some of who have placed orders

Order (OrderID, Custid, orderdate) #has 5000 orders

**I need to find the “number of customers in Customer whose orders appear in the Order table”. How will you modify the following faulty query by adding just ONE word to the SQL? (2 pts)**

# This query gives the answer ‘5000’

select count(c.custid)

from customer c, order o

where c.custid= o.custid;

1. **This question is about Referential Integrity Constraints (RICs). Consider the following schemas. The two tables, Customer and Order are currently empty.**

Customer (CustomerID (PK), Custname)

Order (OrderID (PK), OrderDt, CustomerID (FK))

**I have also created two temporary tables called Table A and Table B and filled them with the following data – they have no RICs.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table A**   |  |  | | --- | --- | | **CustomerID** | **CustName** | | 1 | John Wayne | | 2 | Ingrid Bergman | | 3 | Bruce Lee | | 5 | Omar Sharif | | 6 | Samuel Jackson | | **Table B**   |  |  |  | | --- | --- | --- | | **OrderID** | **OrderDate** | **CustomerID** | | 101 | 01-Jan | 1 | | 102 | 04-Jan | 5 | | 103 | 04-Jan | 6 | | 104 | 05-Jan | 3 | | 105 | 06-Jan | 7 | |

**Using Tables A & B, I would like to populate Customer and Order tables by running the following queries.**

Insert into Customer select \* from A;

Insert into Order select \* from B;

**Explain the result of running each of the above queries.**

**Question 2 (18 points)**

1. Consider the data model for data about donations. **(10 pts)**

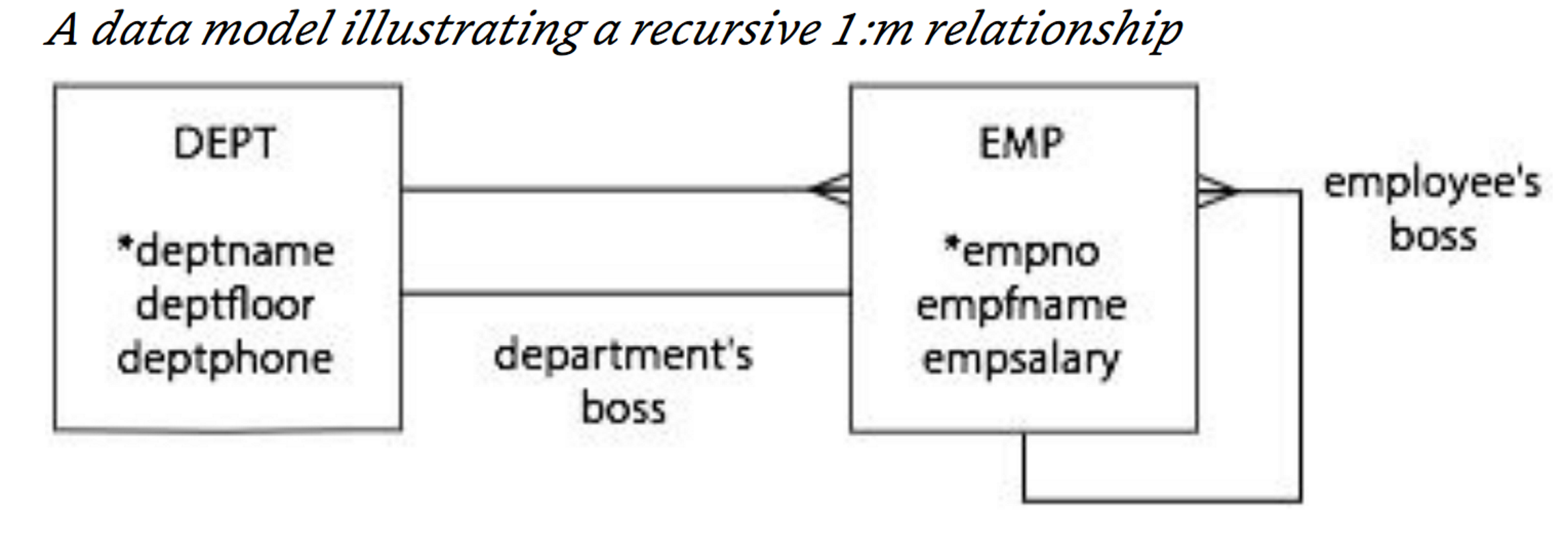
Donor: donorid (PK), dlname, dfname, dphone, dstate, dcity

Gift: donorid (PK, FK), year (PK, FK), amount

Year: year(PK), yeargoal

Create four queries as follows (2.5 pts each):

1. List the donor names who gave more than twice the average.
2. In which years did the total donated exceed the goal for the year?
3. List only those donors (last and first names) who gave more in year 2000 than they did in 1999. (Hint: Either you could use donor table twice with different aliases, or use EXISTS).
4. List the total amount donated by each donor for years 1999 and 2000 combined. The query result should list two columns - the donor names and the total amount.
5. Queries – Recursive (5 points)



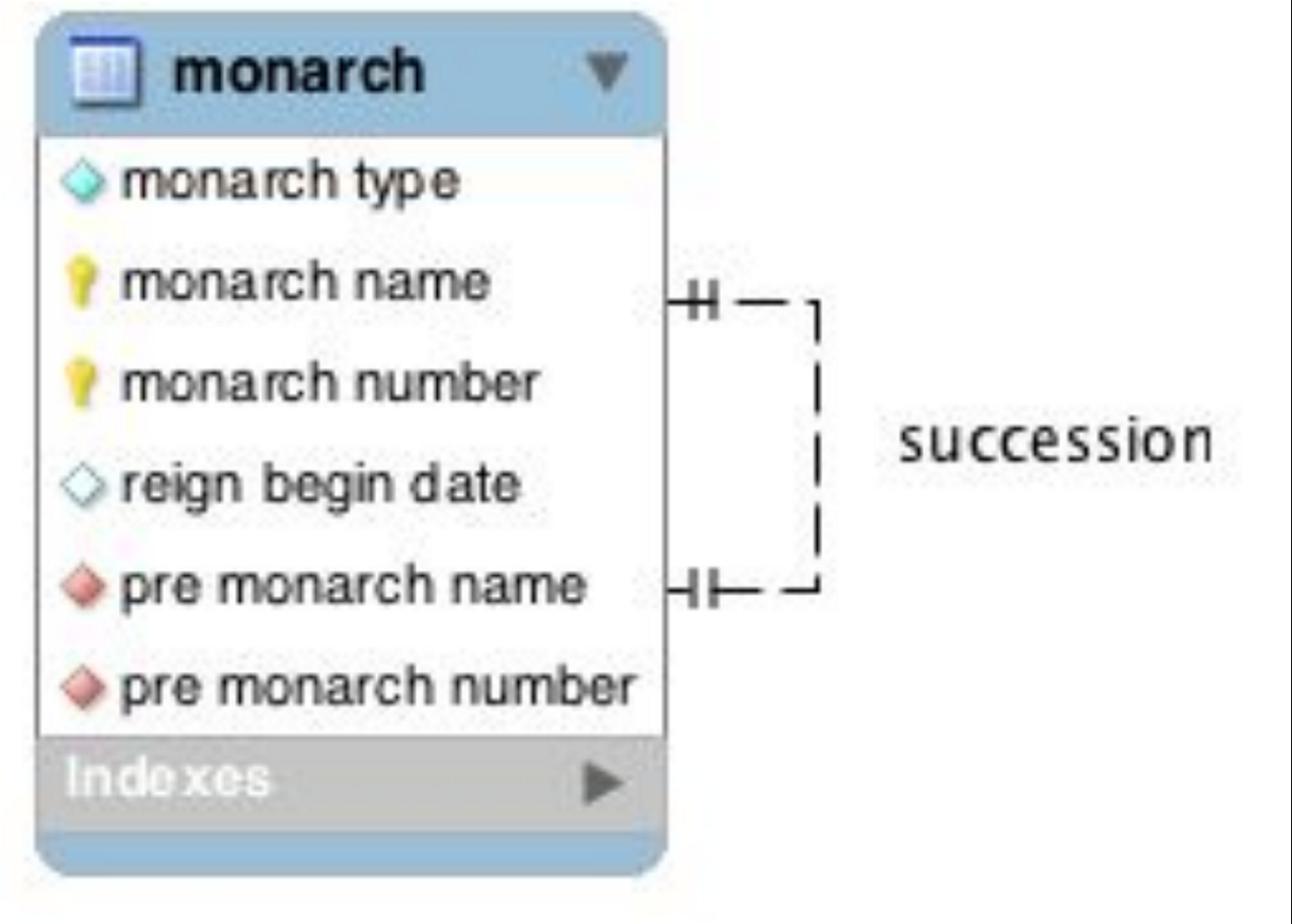
The three RICs are:

Table Dept has empno (FK) as Department’s Boss

Table Emp has deptname (FK) as Department’s name

Table Emp has Bossno (FK) linked to empno (PK) in Table Emp

1. List the departments where the average salary of the employees, **excluding the department’s boss**, is greater than $ 25,000. (2 pts)
2. Find the names of employees who are in the same department as their boss (i.e., the Employee’s boss). (3 pt)
3. Consider the following data model – you may use suitable abbreviation for its fields -- e.g., reign begin date = regbn (3 points)

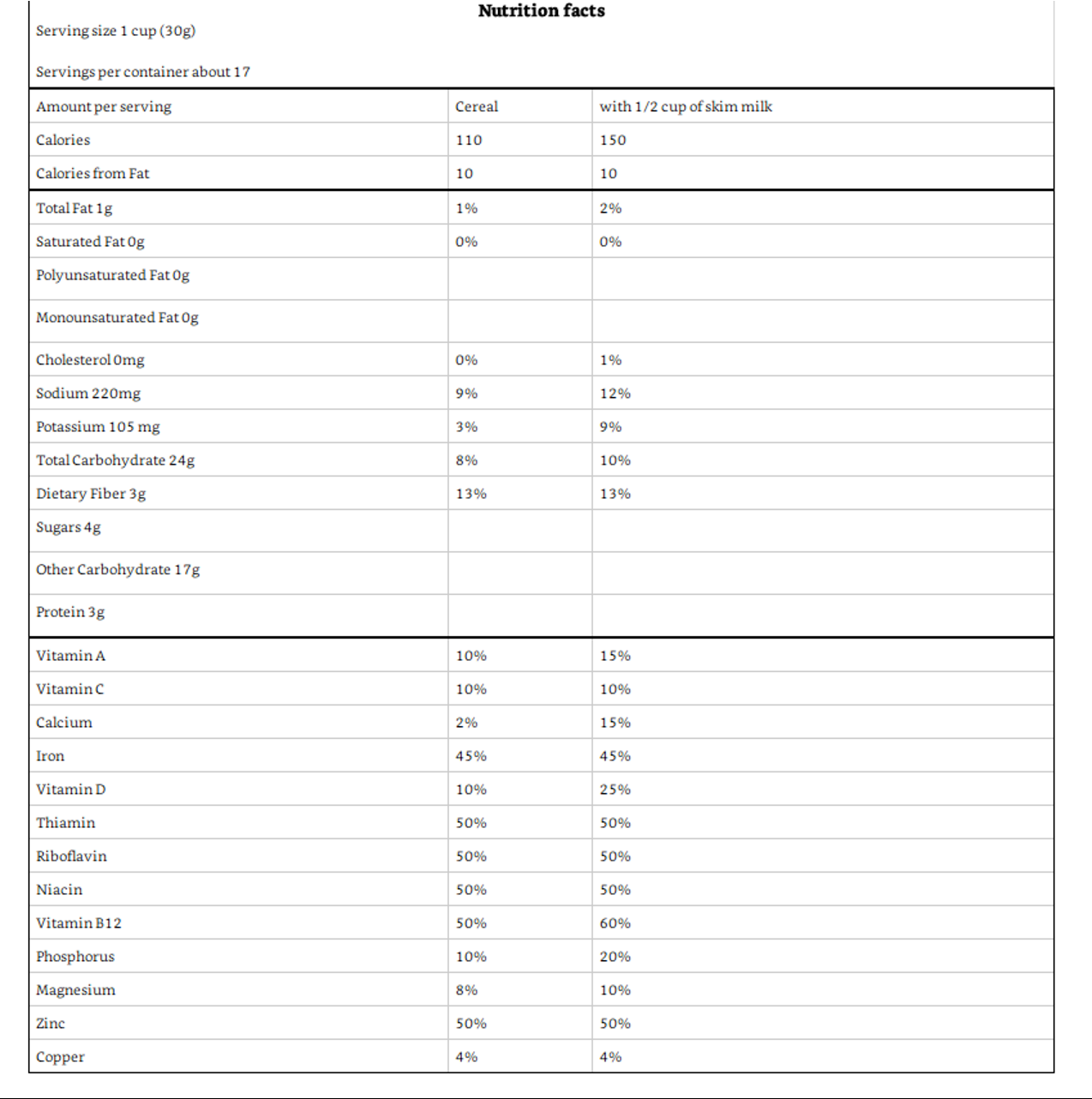


1. How many kings are there in the table? (Hint: monarch\_type (or montype) can be either King or Queen) (1 pts)
2. Which monarch had the shortest reign? (2 pts) Assume reign\_begin date (regbn) is a field of “Date” type. Hint: Find the monarch who has the minimum difference between his/her regbn and the successor’s regbn. The SQL command to find the number of days between any two dates is datediff(date1, date2).

**Question 3 – Data Modelling: Create the physical data models, meaning table schemas that show the entities, attributes, primary keys and foreign keys. You may answer either by listing the table schemas as in Q2a or by drawing the entities and fields as in Q2c. (15 points)**

1. The table in Appendix A records data found on the side of a breakfast cereal carton. Use these data as a guide to develop a data model to record nutrition facts for a meal. In this case, a meal is a cup of cereal and 1/ 2 cup of skim milk. **(5 points)**
2. A travel agency is frequently asked questions about tourist destinations. For example, customers want to know details of the climate for a particular month, the population of the city, and other geographic facts. Sometimes they request the flying time and distance between two cities. The manager has asked you to create a database to maintain these facts. (5 points)
3. Model the facebook data described in the para below. (5 points). A user can invite more than one user to be a friend. A user can have many friends. The gender, name, birthdate and location of each user is captured. A user can create many posts at various date and times. A user can ‘like’ one or more posts of other users. A user can also comment on someone else’s post. For each of the above events, the date-time is always captured (dttime) using which the Facebook data analysts develop statistical models of human behavior.

**Appendix A**



**\_\_\_\_\_\_\_\_EXAM ENDS HERE\_\_\_\_\_\_\_\_**