kasneb

CICT PART II SECTION 3

DATABASE SYSTEMS

WEDNESDAY: 29 November 2017. **Time Allowed: 3 hours.**

Answer ALL questions. Marks allocated to each question are shown at the end of the question.

QUESTION ONE

(a) Citing their goals, differentiate between homogeneous and heterogeneous distributed databases. (4 marks)

(b) Explain four important factors that should be considered in the design of parallel databases. (4 marks)

(c) Distinguish between the following kind of indexes:

(i) "Ordered" and "hash" indexes. (2 marks)

(ii) "Dense" and "sparse" indexes. (2 marks)

(d) Speed-up is one of the optimisation techniques in distributed parallel databases:

Required:

(i) Explain how speed-up in a distributed parallel database is achieved. (2 marks)

(ii) Express the formula for the computation of speed-up in a distributed parallel database. (2 marks)

(e) Distinguish between embedded Structured Query Language (SQL) and Dynamic SQL. (4 marks)

(Total: 20 marks)

QUESTION TWO

Study the tables titled 'VotersRecords' and 'Provinces' below and answer the questions that follow:

VotersRecords

CandidateName	NoOfVotes	Party	ProvinceID
A	20122	Green	P001
B	32333	White	P002
A	20909	Green	P003
A	31789	Green	P002
B	32532	White	P004
B	43332	White	P003
A	2424	Green	P004
B	15212	White	P001

Provinces

ProvinceID	ProvinceName
P001	East
P002	West
P003	North
P004	South
P005	Central

Required:

(a) Write Structured Query Language (SQL) statement for displaying all databases in the DBMS. (1 mark)

(b) The 'Provinces' table, is missing a column by the name 'Country'. Write an SQL statement to modify the table and add a column called 'Country' of datatype 'varchar' and can hold a length of 25 characters. (2 marks)

(c) Write SQL statement that will update all the records in the new column added in (b) above to 'Kenya'. (2 marks)

(d) In the 'provinces' table, central province was added erroneously.
Write an SQL statement that will delete this record. (2 marks)

(e) Write an SQL statement that will display the name of the candidate, the party and the provinces.
The records should be grouped by candidate's name. (3 marks)

(f) Write an SQL statement that will display the name of candidate, sum of votes of each candidate and the name of the party. (3 marks)

(g) Write an SQL statement to display the province that candidate A got the least number of votes. (3 marks)

(h) Highlight four application areas of multimedia databases. (4 marks)

(Total: 20 marks)

QUESTION THREE

(a) Study the following relationship schema:

Supplier (supplier_id:integer, s_name:string, address:string)
Parts (part_id:integer, supplier_id:integer, p_name:string, color: string)
Inventory (inventory_id:integer, part_id:integer, cost: integer)

Required:

(i) Write a query in relational algebra to find names of suppliers who supply red parts. (3 marks)

(ii) Write SQL code from the relational algebra query in (a)(i) above. (4 marks)

(b) As a database administrator, it is important to measure performance of a database. List two tools for doing this job. (2 marks)

(c) Highlight three problems solved in the process of database monitoring. (3 marks)

(d) Describe two functions of mySQL connector used when integrating VB.NET with mySQL database system. (4 marks)

(e) Differentiate between "integrity constraint" and "domain constraint" in the context of database systems. (4 marks)

(Total: 20 marks)

QUESTION FOUR

(a) Distinguish between the terms "simplicity" and "minimality" as used in database design. (4 marks)

(b) (i) Describe an example of functional dependency in database systems. (2 marks)

(ii) Explain transitivity as used in functional dependency. (2 marks)

(c) Citing two examples explain the term "composite attributes" as used in data modelling. (4 marks)

(d) Database security is paramount.

Required:
List four capabilities of the superuser account in database management system (DBMS). (4 marks)

(e) Explain when the lost update problem and dirty read occurs. (4 marks)

(Total: 20 marks)

QUESTION FIVE

(a) Consider the following transactions, T_0 and T_1:

T_0 executes before T_1 and initial values are as follows: A = 100, B = 200 and C = 700.

T_0:	T_1:
read (A)	read (C)
A = A – 50	C = C – 100
write (A)	write (C)
read (B)	
B = B + 50	
write (B)	

Required:

(i) Describe the log-based recovery approach used. (2 marks)

(ii) Extract the corresponding log record after the transactions execute to completion. (6 marks)

(b) Explain the term "data governance" in the context of database systems. (2 marks)

(c) Discuss three scenarios where a view could be helpful in a database management system. (6 marks)

(d) Using an example, describe "basket analysis" in context of database technology. (4 marks)

(Total: 20 marks)

.....