School Name

Grade 12 Information Technology

Normalisation Test 2018

1 hour / 40 Marks

SCENARIO

Architectural company, PlanAhead, has different projects going at the same time. Their staff work on multiple projects concurrently (at the same time) and swap between projects as needed during the week. The company uses a database to keep track of their staff details, project details and the amount of hours each staff member works on a particular project.

All data is stored in a single table named *tblProjects*.

FieldName	**Description**
StaffID	Unique staff identification number
FirstName	Staff member's first name
LastName	Staff member's last name
Phone	Staff member's phone number
Email	Staff member's email address
ProjectID	Unique project identification number
ProjectName	Name of particular building project
Address	Full address of building project
OccupancyClass	A unique code given to a specific type of occupancy
OccupancyType	Type of occupancy i.e. How the building is used
HoursWorked	The amount of hours worked on a particular project

Sample data

Staff ID	First Name	Last Name	Phone	Email	Project ID	Project Name	Address	Occupancy Class	Occupancy Type	Hours Worked
204	Laura	Venter	0823659011	laura@planahead.co.za	201701	House Conradie	14 Willow Lane, Pinelands, Cape Town 7405	RES01	Single Residential	10
98	Mark	Naidoo	0763351089	mnaidoo@telkomsa.net	201701	House Conradie	14 Willow Lane, Pinelands, Cape Town 7405	RES01	Single Residential	7
204	Laura	Venter	0823659011	laura@planahead.co.za	201803	Rose Cottages	14 Outeniqua Road, George 6530	RES09	Cluster Housing	23
133	Lukas	Meyer	0728003544	lucas@planahead.co.za	201803	Rose Cottages	14 Outeniqua Road, George 6530	RES09	Cluster Housing	23
533	Mary	Ntuli	0762752638	mary@planahead.co.za	201702	Redgrave Centre	99 Main Rd, Uniondale, 6460	BUS01	Shopping Centre	5
…	…	…			**…**	…	**…**	…	**…**	…

PlanAhead have encountered some issues whilst using their database so they called in a database design consultant in to investigate. After examining the database she has explained to management that it could be further normalised as it is only in first normal form (1NF).

1. The table has a primary key: identify the key.

   StaffID and ProjectID *Must give both* (1)

2. Explain how *tblProjects* meets the requirements of 1NF.

   It has a primary key ✔

   It has no repeating fields / groups of fields ✔

   It does not have multiple values within a single field ✔ (3)

3. Identify 3 types of anomalies that might occur in the database. Give examples of how each might occur using *tblProjects*.

   Update✔: If an employee changes email / cellphone number, it would have to be changed on multiple records. ✔

   Delete✔: If you delete an employee and they are the only employee assigned to a project, the details of the project will be lost. / If you delete a project and it is the only project an employee is assigned to, the details of the employee will be lost.✔ *Must state the only employee / the only project.*

   Insert✔: You cannot add a new employee if they have not yet been assigned to a project. / You cannot add a new project if no employees have been assigned to it yet. ✔ (6)

4. Apart from preventing anomalies, which other advantage can normalisation provide?

   Saves storage space (because reduces redundant data) / makes it easier to write complex queries ✔ *Do not accept "More efficient" unless they say why.* (1)

5. During the process of normalising the database, the consultant would keep the following concepts in the back of her mind: 2NF, 3NF, primary key and composite key. Write definitions for these concepts:

   5.1 2NF

   A table is in 2NF if it is in 1NF✔ and does not have any partial dependencies✔ (2)

   5.2 3NF

   A table is in 2NF if it is in 2NF✔ and does not have any transitive dependencies✔ (2)

   5.3 Primary key

   A field or group of fields that uniquely identifies✔ records within a table✔ (2)

   5.4 Composite key (2)

A key consisting of more than one field✔✔

6. The consultant has identified a field in the database that contains non-atomic data. She has asked management whether they would like to leave it as is or change it.

   6.1 Define atomic data.

   Data in its simplest form ✔ (1)

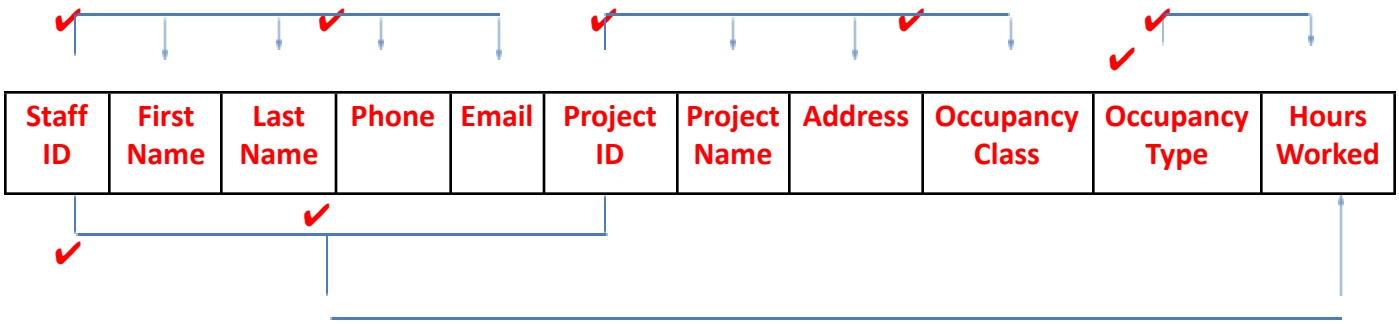
   6.2 Identify an example of non-atomic data in the table.

   Address ✔ (1)

   6.3 Explain how this data might be better stored.

   Could be split into AddressLine1, Suburb, City, Province, PostCode etc ✔ (1)

7. Draw a dependency diagram for *tblProjects*. (7)



8. Normalise the database to 3NF using relational notation.

   tblProjects(ProjectID,✔ ProjectName, Address,✔ OccupancyClass FK)

   tblStaff(StaffID,✔ FirstName, LastName, Phone, Email)✔

   tblOccupancy(OccupancyClass✔, OccupancyType)✔

   tblHoursWorked(ProjectD FK,StaffID✔ FK,HoursWorked)✔ (8)

9. The database consultant mentioned the concept of denormalisation. She explained it as follows:

   "Denormalisation is a strategy used on a previously-normalized database to increase performance. In computing, denormalisation is the process of trying to improve the read performance of a database, at the expense of losing some write performance, by adding redundant copies of data or by grouping data. It is often motivated by performance or scalability in relational database software needing to carry out very large numbers of read operations. Denormalisation should not be confused with Unnormalised form. Databases/tables must first be normalized to efficiently denormalise them."

   (https://en.wikipedia.org/wiki/Denormalization)

9.1 Why do you think write performance is compromised in a denormalised database?

Update anomalies will cause data to be written multiple times when changed, instead of just once in a normalised database. *Any good answer indicating understanding* (1)

9.2 Do you think it would be recommended for *PlanAhead's* database to be denormalised? Explain your answer using information from the paragraph above.

No, the reason for denormalisation is to improve read performance for very large databases.✔ PlanAhead will not have a very large database✔ (like Amazon, Google etc) (2)

[40]