IT3 Database Questions

1

15	Explain what is meant by a relational database. [2]
A (large) collection of data items and links/relationships between them (1) structured in such a way that it allows it to be accessed by a number of different applications programs (1). **OR A group of tables linked together (1) by primary and foreign keys (1).**	
16	A car hire company uses a relational database management system for organising its rentals. For simplicity, a person may only hire one car at a time. Part of one table in the car hire database could be: CAR [CarID, Make, Model, Colour, Year of registration, Hire rate, ...] with CarID being the primary key. Give two other suitable tables you could expect to see in this database, identifying any primary and foreign keys and at least two other fields in each table. [7]
CUSTOMER [CustomerID, Title, Firstname, Surname, Address1 …..] **HIRE [HireID, CarID*, CustomerID*, DateOut, ReturnDate]** **Where underscore = primary key and *=foreign key** **1 mark per table name = 2** **1 mark per primary key = 2 (one of the marks could be for a composite field in HIRE** **1 mark each for two foreign keys = 2 (If**	
17	Describe how the car hire database structure could be improved to allow a person to hire more than one car at a time. [1]
Either: Add a link table/form. **Make it a one to many relationship**	
18	Using appropriate examples, describe the advantages of a relational database approach over a flat file approach. [8]

(Security) Hierarchy of passwords limits users to various parts of the program.(1) A receptionist would only have access to basic customer details whilst a manager would see all information on the customer account.(1).

(Security) Access rights to parts of the program only certain users can access and change data.(1) A clerk would see all the information on a customer account but be unable to alter the hire rate details whilst a manager could.(1).

Consistency - Data consistency is the relationship between the input data, the processed data and the output data as well as other related data.(1) If the system is working properly the data will be correct at each stage and is said to be consistent.(1)
OR
Data consistency is using one file to hold a central pool of data. / A company may hold all its customer data in one file. (1) This avoids the need to input data twice so that if data is changed in one file it won't need to be changed in another and remains consistent.(1).
OR
Data being inconsistent in a flat file due to possibility of different formats etc (1) and being consistent in a RDBMS as each record is only stored once so cannot have different attributes(1).

Redundancy Data redundancy is where you store an item of data more than once / A company may hold its data in different files.(1) This is wasteful because some data may need to be input twice and if data is changed in one it will need to be changed in the other. / Data which is repeated unnecessarily is called redundant data.(1).

Independence Data independence – the data and the applications/programs used to access it are independent/separate.(1) New applications can be developed to access the data without changing the data / New systems can still use existing data. (1).

15	A Hospital uses a relational database for storing patient records. Explain what is meant by a relational database. [2]

A (large) collection of data items and links between them (1) structured in such a way that it allows it to be accessed by a number of different applications programs (1)
A group of tables linked (1) together by primary and foreign keys (1)

16	Data consistency, Data redundancy, Data integrity and Data independence are all terms associated with relational databases. Explain what each of these four terms means. [4]

Data consistency is where the same data has exactly the same value and formatting throughout, (normally because it is only stored once)
OR
Data consistency is the relationship between the input data, the processed data and the output data as well as other related data. If the system is working properly the data will be correct at each stage and is said to be consistent.
OR
Data consistency is using one file to hold a central pool of data this means if information is changed in one file it does not need to be changed in another.
OR
Data being inconsistent in a flat file due to possibility of different formats etc and being consistent in a RDBMS as each record is only stored once so cannot have different attributes.
Data which is repeated unnecessarily is called redundant data.
OR
Data redundancy is where you store an item of data more than once.
Data integrity – the correctness of the data. Truthful, Accuracy.
Data independence – the data and the applications/programs used to access it are independent/separate.
OR
New applications can be developed to access the data without changing the data.

17	Consultants and patients are allocated to wards but each patient has only one consultant and each ward has their own consultant. One table in this database could be: WARD(Wardid, WardName, NumOfBeds) With Wardid being the primary key. Give two other suitable tables you could expect to see in this database, identifying any primary or foreign key. [6]

PATIENT (Patientid, surname, phone, Wardid#, Consultantid#)
CONSULTANT (Consultantid, surname, phone, etc)
Underline = primary, # = foreign
1 mark per table name
1 mark per foreign key
1 mark per primary key

18	Describe what a data warehouse is and give an example of how a company might use it. [3]

Large, Archive and used for Decision Making – Look for 2 of these 3
Up to two marks for an example of use
Examples for one mark (What or Why)
Allows the company to store information about every sale. (1)
Allows the company to see who has bought what items and when. (1)
Can use it to plan future changes or developments in their business. (1)
Allows the company to use data mining. (1)
Allows the company to find the most popular product (1)

19	Explain what is meant by data mining and give a detailed example of how a company might use it. [3]

One mark for the meaning (patterns / trends / generating new information)
Data mining is interrogating the data to find patterns in the data which is stored in the warehouse.
Examples worth 1 mark:
Can provide:
☐ **The company with a list of customers likely to buy a certain product, which they can then use to target with a mail shot.**
☐ **Comparisons with competitors**
☐ **Useful 'what if' results from modelling exercises**
☐ **Predictions for future sales**
☐ **Analysis of best sites for shops**
☐ **Analysis of sales patterns**
☐ **Returned information can be tested for plausibility.**
☐ **Data if of value can be processed into a report to help decision making.**

15	Explain what is meant by a relational database. [2]

A (large) collection of data items and links/relationships between them (1) structured in such a way that it allows it to be accessed by a number of different applications programs (1)
OR
A group of tables linked together (1) by primary and foreign keys (1).

16	A DVD Rental Company uses a relational database management system for organising its lending. For simplicity, a person may only borrow one DVD at a time. Part of one table in the DVD Rental Company database could be: BORROWER [BorrowerID, FirstName, Surname, Borrower Type, Address1, ...] with BorrowerID being the primary key. Give two other suitable tables you could expect to see in this database, identifying any primary and foreign keys. [7]

DVD [DVDID, Title, Genre, Classification, …..]
LOAN [LoanID, DVDID*, BorrowerID*, DateOut, ReturnDate]
Where underscore = primary key and *=foreign key
1 mark per table name = 2
1 mark per primary key = 2
1 mark for two foreign keys = 1 (If key duplicated - no mark)
1 mark per table for 2 extra fields = 2

17	Relational databases are normalised to work effectively. Explain what is meant by the term data normalisation. [2]

A staged (mathematical) process (1) which removes repeated groups of data and inconsistencies. (1)
Or
Simplifying data structures (1) so that attributes in each table only relate to the entity. (1)
Marks can be gained by using an example of the process of going from first to third form.
Do NOT accept advantages of databases

18	Some companies now use distributed databases. Explain what is meant by the term distributed databases and explain their advantages and disadvantages. [6]

A distributed database is a single database that is under the control of a DBMS where the storage devices are not all attached to a common processor (1). Instead the data is stored in storage devices attached to multiple computers usually located across a network (1).
Or
A distributed database has data stored on a number of computers at different locations (1) but appears as one logical database (1).
Advantages
☐ **If data lost on central site it could be reduplicated from local site**
☐ **Allows sharing of the data and the results of processing the data**
☐ **New locations can be added to the database without the need for rewriting the entire database**
☐ **Faster response to user queries of the database**
☐ **Non-dependence on one central huge store of data**
☐ **Easy to backup and copy data from one server to another**
☐ **If one server fails then the other servers can be used**
☐ **Reduces network traffic as local queries can be performed using the data on the local server.**
Disadvantages
☐ **Software more complex than a centralised database system**
☐ **If data is transferred it presents more of a security risk from hackers**
☐ **As all the data is not stored in one location if a local site does not have adequate backup then this data might be lost to others**

If data is stored and updated in more than one place there is an increased chance of data inconsistency

 Heavy reliance on networks and communications which may not always be reliable

 Security issues particularly if sensitive personal data is being transferred

 If one of the links to a server failed then the data could not be obtained from that server

 Increased costs owing to the use of expensive communication lines. NOT just costs.

13	Relational databases are normalised to work effectively. Explain what is meant by the term data normalisation. [2]

A staged (mathematical) process (1) which removes repeated groups of data and inconsistencies. (1)
Or
Simplifying data structures (1) so that attributes in each table only relate to the entity. (1)
Or
Normalisation is the organisation

14	Give three reasons why relational databases are more secure than a flat file approach. [3]

Hierarchy of passwords -- passwords to see separate parts
NOT just passwords
Storage of data separate to programs
Access rights to parts of the program.

15	Distributed databases are widely used. Explain what is meant by the term distributed databases and explain their advantages and disadvantages. [7]

A distributed database is a single database that is under the control of a DBMS where the storage devices are not all attached to a common processor (1). Instead the data is stored in storage devices attached to multiple computers usually located across a network (1).
Or
A distributed database has data stored on a number of computers at different locations (1) but appears as one logical database (1).
1 mark each for any four points but must have at least one advantage and one disadvantage
Advantages
☐ If data lost on central site it could be reduplicated from local site.
☐ Allows sharing of the data and the results of processing the data.
☐ New locations can be added to the database without the need for rewriting the entire database.
☐ Faster response to user queries of the database.
☐ Non-dependence on one central huge store of data.
☐ Easy to backup and copy data from one server to another.
☐ If one server fails then the other servers can be used.
☐ Reduces network traffic as local queries can be performed using the data on the local server.
Disadvantages
☐ Software more complex than a centralised database system.
☐ If data is transferred it presents more of a security risk from hackers.
☐ As all the data is not stored in one location if a local site does not have adequate backup then this data might be lost to others.
☐ If data is stored and updated in more than one place there is an increased chance of data inconsistency.
☐ Heavy reliance on networks and communications which may not always be reliable.
☐ Security issues particularly if sensitive personal data is being transferred.
☐ If one of the links to a server failed then the data could not be obtained from that server.
☐ Increased

16	Explain what is meant by a data warehouse and describe the advantages it might give a retail company. [3]

Large, Archive and used for Decision Making – Look for 2 of these 3
A large collection of archived data used for decision making (1)
OR
A large company generates huge quantities of data stored in a consistent order to make interrogation more productive.(1)

OR

Data is non-volatile and time invariant (archive data).Used to support organisational decision making.(1)

OR

A huge database specifically structured for information access and reporting (1)

Up to two marks for an example of use

Examples for one mark (What or Why)

Allows the company to store information about every sale. (1)

Allows the company to see who has bought what items and when. (1)

Can use it to plan future changes or developments in their business. (1)

Allows the company to use data mining. (1)

Allows the company to find the most popular product. (1)

Example for two marks (What and Why)

Allows the company to see who has bought what items (1) and then target them with special offers. (1) (why)

17	Explain what is meant by data mining and describe an example of how a retail company might use it. [3]

Data mining is interrogating the data to find patterns in the data which is stored in the warehouse.

Examples worth 1 mark:

Can provide:

☐ **The company with a list of customers likely to buy a certain product, which they can then use to target with a mail shot.**

☐ **Comparisons with competitors.**

☐ **Useful 'what if' results from modelling exercises.**

☐ **Predictions for future sales.**

☐ **Analysis of best sites for shops.**

☐ **Analysis of sales patterns.**

☐ **Returned information can be tested for plausibility.**

☐ **Data if of value can be processed into a report to help decision making.**

17	Data in a relational database is normalised. Explain what is meant by a relational database and data normalisation. [4]

A large collection of data items and links between them (1)
(structured in such a way that) it allows it to be accessed by a number of different
applications programs (1) / a group of tables linked (1) together by primary and foreign
keys (1).
Suitable definition of data normalisation, such as:
A staged (mathematical) process (1) which removes repeated groups of data and
inconsistencies. (1)
Or
Simplifying data structures (1) so that attributes in each table only relate to the entity. (1)
Or
Normalisation is the organisation of data into tables (1) which relate to a single entity. (1)

18	Describe the terms data redundancy, data integrity and data consistency in relation to a relational database approach over a flat-file approach. [3×2]

Data redundancy
is where you store an item of data more than once.
A company may hold its data in different files. This is wasteful because some data may
need to be input twice. Data which is repeated unnecessarily is called redundant data.
It refers to the unnecessary duplication of data.
In a flat-file database details about such information as customer details will be
duplicated. In a well designed relational database there should be no 'repeating
attributes', no piece of data should be unnecessarily repeated.
Data integrity
The integrity of data is the correctness, i.e. the extent to which it truthfully represents the
original information. Truthful, Accuracy.
One of the problems of maintaining integrity arises when updating occurs, and every
record has to be changed in a flat-file database, if one record was left unchanged the
data would no longer be wholly correct. In a relational database you only have to change
data in one table and all other references in any other application will automatically be
changed.
Data consistency
Data consistency is the relationship between the input data, the processed data and the
output data as well as other related data. If the system is working properly the data will be
correct at each stage and is said to be consistent.
OR
Data consistency is using one file to hold a central pool of data. A company may hold all
its customer data in one file. This avoids the need to input data twice so that if data is
changed in one file it won't need to be changed in another and remains consistent.
OR
Data being inconsistent in a flat file due to possibility of different formats, etc, and being
consistent in a RDBMS as each record is only stored once so cannot have different
attributes.
A date field could be stored in file as a text field but in another field as a date/time field
and the data would be incompatible. In a relational database because the attributes of
any one entity are contained within one file, there is no risk of the same attribute being
stored in a different format in a different file (Spelling mistakes in names).

19	Give two reasons why relational databases are more secure than the flat-file approach. [2]

Any two from:
Hierarchy of passwords - passwords to see separate parts
NOT just passwords
Storage of data separate to programs

Access rights to parts of the program.	
20	Modern technology also makes use of distributed databases. Define what is meant by a distributed database. Describe two security issues that apply to distributed databases and suggest a different method for overcoming each issue. [5]

A distributed database is a single database that is under the control of a DBMS where the storage devices are not all attached to a common processor. Instead the data is stored in storage devices attached to multiple computers usually located across a network.
OR
A distributed database has data stored on a number of computers at different locations but appears as one logical database.
Computers are located on a number of sites so it is important to ensure only authorised users can access the system.
Can be achieved by using passwords to authorised users, and regularly updating the passwords to increase the levels of security.
Data regularly transmitted between different sites and so data may become corrupt or be tampered with during transmission.
Checks are put in place in order to ensure that the data that arrives is both secure and accurate.
Encryption of transmitted data.
Viruses if linked to a local database
Use a firewall and anti-virus software
Problem hacking into local data Hacking into transmitted data are separate problems

17	Many companies would now struggle to operate effectively without database technology. Explain what is meant by a relational database. [2]

A relational database is a large collection of data items/tables and links between them (1 mark)

Plus either of the following for second mark
• structured in such a way that allows it to be accessed by a number of different applications programs
• use of primary and foreign keys.

18	Using appropriate examples, describe the advantages of a relational database approach over a flat file approach. [6]

Hierarchy of passwords **limits users to various parts of the program. A receptionist would only have access to basic patient details whilst a doctor would see all information on the patient.**

Access rights **to parts of the program only certain users can access and change data. A nurse would see all the information on a patient but be unable to alter the drug dose details whilst a doctor could.**

Consistency **- Data consistency is the relationship between the input data, the processed data and the output data as well as other related data. If the system is working properly the data will be correct at each stage and is said to be consistent.**
OR
Data consistency is using one file to hold a central pool of data. A company may hold all its customer data in one file. This avoids the need to input data twice so that if data is changed in one file it won't need to be changed in another and remains consistent.
OR
Data being inconsistent in a flat file due to possibility of different formats etc and being consistent in a RDBMS as each record is only stored once so cannot have different attributes

Redundancy **Data redundancy is where you store an item of data more than once / A company may hold its data in different files. This is wasteful because some data may need to be input twice and if data is changed in one it will need to be changed in the other. / Data which is repeated unnecessarily is called redundant data.**

Independence **Data independence – the data and the applications/programs used to access it are independent/separate. / New applications can be developed to access the data without changing the data./**
New systems can still use existing data.

19	Describe what a data warehouse is and give an example of how a company might use it. [3]

Large, Archive and used for Decision Making – Look for 2 of these 3
A large company generates huge quantities of data stored in a consistent order to make interrogation more
productive.(1)
Data is non-volatile and time invariant (archive data).Used to support organisational decision making.(1)
Up to two marks for an example of use
Examples for one mark
Allows the company to store information about every sale. (1)
Allows the company to see who has bought what items and when. (1)
Can use it to plan future changes or developments in their business. (1)
Allows the company to use data mining. (1)
Allows the company to find the most popular product and buy more (1)
Example for two marks
Allows the company to see who has bought what items (and why) and then target them

	with special offers
20	Explain what is meant by data mining and give a detailed example of how a company might use it. [3]

Data mining is interrogating the data to find patterns in the data which is stored in the warehouse.
Up to two marks for example of use
Look for the idea of trends, patterns or generating new information
• is the analysis of a large amount of data in a data warehouse to provide new information.
• Is interrogating large amounts of data
• is a speculative process investigating potential patterns
• involves the presumption that dormant within the data are undiscovered patterns / groupings / sequences / associations.
• software uses complex algorithms to search for patterns.
• is drilling down into the mass of data so users can understand it more / discover meaningful patterns.
• Is looking for meaningful patterns in a large mass of data and presenting results in tables and graphs.
Examples worth 1 mark:
Can provide:
• the company with a list of customers likely to buy a certain product, which they can then use to target
with a mail shot.
• comparisons with competitors
• useful 'what if' results from modelling exercises
• predictions for future sales
• analysis of best sites for shops
• analysis of sales patterns
• returned information can be tested for plausibility.
• data if of value can be processed into a report to help decision making.

21	Discuss in general terms, the advantages and disadvantages distributed databases give to a company. [3]

Must be in context of a company
Advantages
• If data lost on central site it could be reduplicated from local site
• Allows sharing of the data and the results of processing the data
• New locations can be added to the database without the need for rewriting the entire database
• Faster response to user queries of the database
• Non-dependence on one central huge store of data
• Easy to backup and copy data from one server to another
• If one server fails then the other servers can be used
• Reduces network traffic as local queries can be performed using the data on the local server.
Disadvantages
• Software more complex than a centralised database system
• If data is transferred it presents more of a security risk from hackers
• As all the data is not stored in one location if a local site does not have adequate backup
then this data might be lost to others.
• If data is stored and updated in more than one place there is an increased chance of data
inconsistency.
• Heavy reliance on networks and communications which may not always be reliable

- Security issues particularly if sensitive personal data is being transferred
- If one of the links to a server failed then the data could not be obtained from that server
- Increased costs owing to the use of expensive communication lines.