

Computer Science - Core Skills

Semester 1

BSC109224 Group A

Assignment 6

Student: Lucas Madeira Maranhão

Student ID: 76990

Lecturer: Bernard Joseph Roche

SUMMARY

| | |
|---|---|
| SUMMARY | 2 |
| 1. Data Management Strategies..... | 3 |
| 1. Data collection:..... | 3 |
| 2. Data storage:..... | 3 |
| 3. Data security: | 3 |
| 4. Data analysis:..... | 3 |
| 5. Data governance: | 3 |
| 6. Scalability: | 3 |
| 1. A List of People to Whom You Plan to Send Christmas Cards..... | 4 |
| Storage Methods: | 4 |
| Scalability: | 4 |
| 2. A Database of Books and Their Borrowers in Dorset College's Library | 5 |
| Data Structure: | 5 |
| Storage Methods: | 5 |
| Retrieval Efficiency:..... | 5 |
| Scalability: | 5 |
| 3. All Phone Numbers in the USA, Including Both Landline and Mobile Numbers | 5 |
| Data Structure: | 5 |
| Storage Methods: | 6 |
| Retrieval Efficiency:..... | 6 |
| Scalability: | 6 |
| Conclusion | 6 |
| Question 2 (50%): Basic SQL | 7 |
| PART A: | 7 |
| REFERENCES:..... | 8 |

1. Data Management Strategies

It is a practice way that organizations use the system to collect, protect, store, retrieve and analyse data. This kind of strategy are very important to guarantee the data security, accessibility, integrity and the efficiency during the life. Before going to the given list, I will talk about some key components of effective data management that are important:

1. Data collection:

Identify information that must be collected, based on business goals, including, system, user inputs and data providers.

2. Data storage:

Pick up the appropriated storage solution, like databases or cloud storage.

3. Data security:

Restrict access to sensitive data, requiring authorization and authentication.

4. Data analysis:

Getting data from multiple searches to give a better understanding of analysis.

5. Data governance:

Clear policies of data usage, quality.

Assign responsibilities for data management to get responsibility and oversight.

6. Scalability:

Plan that can grow according to user demands.

New technologies or data sources integrated.

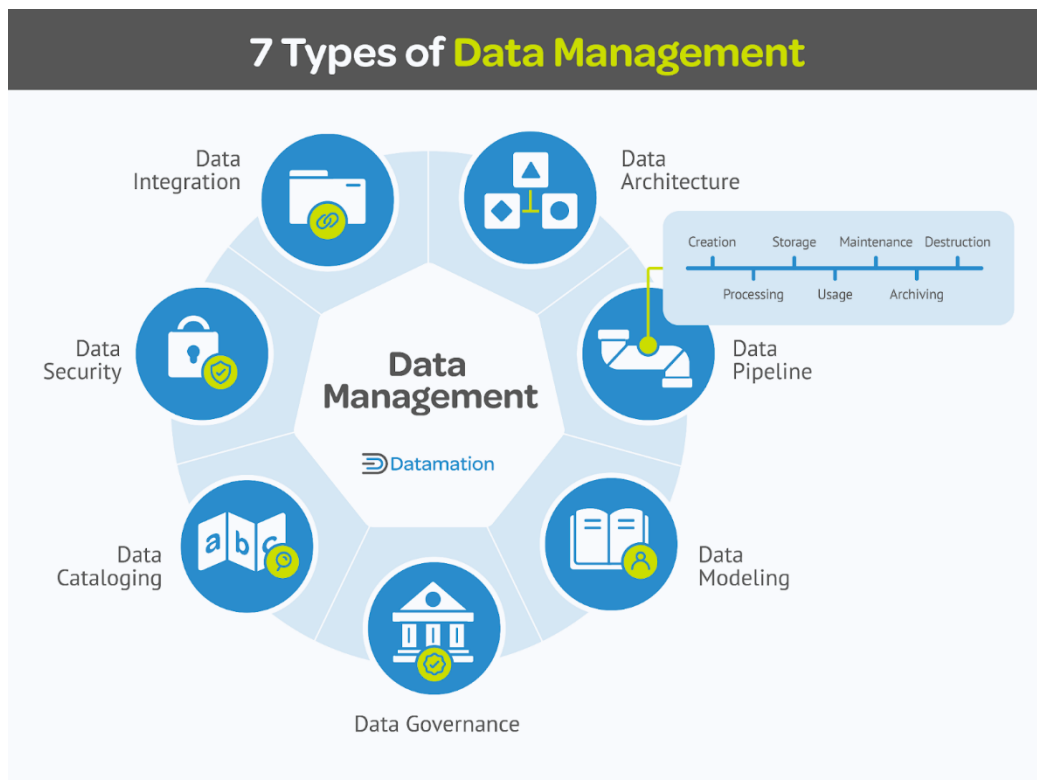


Figure 1 - data management

1. A List of People to Whom You Plan to Send Christmas Cards

Data Structure:

For a Christmas card list, the data structure can be simple. It can be used a piece of paper or use notepad to create a list including Name, Address, Email, phone number, relationship (family, friend).

Storage Methods:

Writing in a piece of paper or a notepad is efficient because is not a complex situation, and the size of people are short.

Retrieval Efficiency(security):

Retrieval efficiency can be saved in your computer, if you wrote in a notepad or word, put in a flash drive or even thou put in cloud storage depending on how important is to you.

Scalability:

It will depend on how many people you want to send the Christmas card, if the quantity of people grows, you can move to a spreadsheet on excel.

2. A Database of Books and Their Borrowers in Dorset College's Library

Data Structure:

to handle with a library database, a related spreadsheet on excel structure is appropriate. It's necessary creating some tables: for Books and Borrowers. The Books table can include a list, which have BookID, Title, Author, Genre, and Year of publication, while the Borrowers table could include Borrower ID, Name, Contact Information, and borrowed Book ID.

Storage Methods:

The better way of storage in this situation is using spreadsheets on excel.

Retrieval Efficiency:

The retrieval efficiency supports detailed queries. Indexing is supported to facilitate fast searching.

Scalability:

A relational database works very well to keep an increasing number of books and borrowers. As the library grows, additional tables can be added (for example: genres, categories, or reservations), and the database can handle larger datasets without significant performance degradation. Regular database maintenance (like checking stock or archiving old data) also helps maintain efficiency as the volume increases.

3. All Phone Numbers in the USA, Including Both Landline and Mobile Numbers

Data Structure:

The data structure for managing phone numbers can be complex because of the vast volume of data. A relational database can again be beneficial, with a single table for phone numbers that includes fields like phone ID, phone number, owner name, Type (Landline or Mobile) and Region.

Storage Methods:

having a big scale (millions of entries), a robust RDBMS is necessary. Solutions like Amazon RDS or Google Cloud SQL can give the infrastructure needed to handle such large data. Additionally, using NoSQL databases might be considered for flexibility, especially if the data structure needs to evolve over time.

Retrieval Efficiency:

For efficient retrieval, adding indexing on the phone number field and possibly the owner's name field would turn into better search capabilities. Additionally, creating stored procedures for common queries for example: retrieving all numbers by region can improve performance. Given the potential size of the dataset, employing sharding (separate the database across multiple servers) could increase the retrieval efficiency so much.

Scalability:

The choice of a scalable storage solution is critical here. Cloud-based databases can automatically scale resources based on demand, accommodating growth as more numbers are added or as user queries increase. Implementing data partitioning and archiving older records can help manage data size effectively while maintaining performance.

Conclusion

In conclusion, data management strategies can have 3 different ways to manage and create, it will depend on the context, size and the complexity of the data. For a Christmas card list, just a piece of paper is enough, while a spreadsheet is essential to managing a library's books and borrowers. On the other hand, managing all phone numbers in the USA requires a vast and scalable database solution. Understanding the differences and choosing appropriate data structures storage methods, and retrieval strategies are crucial for a good data management nowadays.

Question 2 (50%): Basic SQL

PART A:

1 – SELECT * FROM Books.

2 – SELECT Title, Author FROM Books WHERE Genre = "Fiction"

3 – SELECT Title, Author FROM Books WHERE YearPublished > 2000.

4 – SELECT Title, Author FROM Books WHERE Genre = "Fiction" AND YearPublished > 2000

5 – SELECT Genre, COUNT(*) AS NumberOfBooks FROM Books GROUP BY Genre

6 – SELECT * FROM Books WHERE Author = "J.K. Rowling"

7 – SELECT Title, YearPublished FROM Books ORDER BY YearPublished DESC

8 – SELECT Title, Author FROM Books WHERE Title LIKE "Magic"

9 – SELECT Title, YearPublished FROM Books ORDER BY YearPublished ASC LIMIT 1

10 – SELECT Author, COUNT(*) AS TotalBooks FROM Books GROUP BY Author

11 – SELECT Title, Author, YearPublished From Books WHERE YearPublished > "1990" AND YearPublished <= "2010" ORDER BY YearPublished ASC

REFERENCES:

<https://www.actian.com/what-is-a-data-management-strategy/#:~:text=Data%20Management%20Strategy%20is%20enabled,the%20organization's%20other%20strategic%20projects>.

<https://libguides.ucd.ie/data/dmp>

<https://www.datamation.com/big-data/what-is-data-management/>