ITMD 523	*Advanced Database Management*	Homework 1
Student Name	**Harshal Sawant**	Due Date 09/09/23
Instructor	**Prof. Luke Papademas**	Section 05

Part	1	2	3	4	Total
maximum	**25** points	**25** points	**25** points	**25** points	**100** points
Your Score					

Topic: Database Systems

Reading Assignment: Thoroughly read Chapter 1 in the course textbook(s).

Part 1 Glossary Terms - Compare and Contrast

Review, in detail, each of these glossary terms from the realm of computer database systems and computer topics, in general. Then, in your own words, compare and contrast the terms. If applicable, use examples to support your definitions. Consult your notes or course textbook(s) as references.

(a) Structural Dependence VERSUS Structural Independence

Structural Dependency:

In a database system, structural dependency occurs when the format or structure of the data is tightly related to how it is stored or structured. At the end of the day, modifications to the data set's design might significantly affect how clients access or use information. Consequently, the structure becomes dependent on the data, limiting its adaptability and making it more challenging to meet shifting requirements.

Example: Consider a database that contains data on goods and is divided into sections like "Electronics," "Clothing," and "Accessories ". The database structure may need to be changed if you wish to add a new category, such as "Vehicles," which may change how current data is accessed and shown.

Structural Independence:

Structural independence has independent data of database regardless of it's structure. Changes to the database's structure in these circumstances have no impact on how data is stored, processed, or retrieved. It enables greater adaptability and flexibility since you may change the database structure without affecting the data originally present.

Example: For structural independence, you can easily add the "Vehicles" category to the same product database without altering how current information on "Electronics", "apparel", or "Accessories" is arranged or accessible. Changes to the structure of data have no impact on it.

(b) Problems with File System Data Processing VERSUS Problems with Database Management Systems

The problems are as follows:

1. Lengthy development times

File System: Even simple data retrieval activities in a file system need substantial programming. In programming, both what must be done and how it is done must be specified.

DBMS: Present databases feature a non-procedural language for modifying data that

enables users to define what must be done without digging into the details of how to do it. As a result, development times are minimized, and data retrieval is boosted.

2. Difficulty of getting quick answers

File System: Because programmers must create programs for even the most basic reports, file systems make ad hoc inquiries virtually difficult and cause delays in information retrieval.

DBMS: Ad hoc querying is possible with DBMS, allowing users to ask inquiries and instantly obtain data. This makes accessing data fast and versatile.

3. Complex system administration

File System: Managing files in the system gets more difficult as the amount of them rises. Every file needs its own management software.

DBMS: DBMS simplifies data management, improving the effectiveness of system administration. It cuts out the requirement for different file management solutions for every data collection.

4. Lack of security and limited data sharing

File System: File systems often do not have reliable security measures and are not suitable for data sharing among users who are spread out geographically.

DBMS: Contemporary DBMS include strong security features including encryption, access controls, and password protection, making it safer for data transfer between authorized users.

5. Extensive programming

File System: File systems often do not have reliable security measures and are not suitable for data sharing among users who are spread out geographically.

DBMS: Contemporary DBMS include strong security features including encryption, access controls, and password protection, making it safer for data transfer between authorized users.

(c) Data Dependence VERSUS Data Independence

Data dependence:

Data dependence can be defined as the condition in which the modification and accessing of the data is dependent on the data storage characteristics. This means that, if the data storage characteristics are altered then it can affect the program's ability to access the data.

Data independence:

Data independence exists when the modification and accessing the data/information is unaffected by the changes at the data storage characteristics.

Hence, it is simpler to manage software systems when data is independent because you can alter how data is stored without having to rewrite numerous programs. Furthermore, it increases the adaptability of systems, enabling them to accommodate evolving technological demands and data requirements, making it more flexible, unlike the brittle data dependent systems. However, when a program is data dependent, it becomes more difficult since it must understand every aspect of data storage. In summary, data independence makes databases and files easier to maintain as it makes systems more adaptable, but data dependency might be more difficult since it glues

data and programs together.

(d) Data Redundancy VERSUS Data Anomalies

Data redundancy:
The process of storing identical data in several places or files, resulting in multiple copies of the same information, is known as data redundancy. It takes place when identical data is mistakenly replicated across many locations or files.
Data Anomalies:
Data anomalies arise when redundant data forces modifications to be performed in various locations. It occurs because of the requirement to preserve consistency if duplicate data is updated, added, or removed.

In database management, data redundancy and data anomalies are concepts that are closely connected. Data anomalies are problems with consistency when working with duplicate data, whereas data redundancy is the excessive duplication of data across many places or files. Data redundancy causes problems with data integrity, inconsistent data, inadequate data security, and data entry mistakes. In contrast, data anomalies refer primarily to issues that arise when duplicate data must be changed in many locations. These issues include update anomalies, insertion anomalies, and deletion anomalies. As noted in the accompanying text, these concerns highlight how very important proper database architecture and maintenance are to guarantee data correctness, consistency, and security.

(e) Database Designers VERSUS Database Developers

Database Designers:
Database designers lay up the blueprint for how a database should function. Even highly skilled programmers and database specialists cannot create an effective database if they perform an unsatisfactory job. The position of the database designer has expanded due to the desire of businesses to make the best use of their data. SQL, DBMS, data modelling, normalization, data security, and performance optimization are all necessary abilities for database designers. Data analysis, problem-solving, and effective communication are crucial for converting business requirements into effective schemas. Success in complicated projects also depends on documentation, remaining current, and adopting project management and quality assurance procedures.

Database developers:
Database developers are professionals who oversee creating, deploying, and managing databases that safely and effectively handle and store the data of a business. By building and optimizing database structures, crafting challenging SQL queries, and assuring data integrity and efficiency, they play a crucial part in the software development process. Data analysts and software engineers collaborate closely with database developers to understand the data demands of applications and make sure that data is accessible and arranged for a variety of business objectives. Strong analytical abilities, expertise with database management systems like SQL Server, Oracle, or MySQL, and a thorough grasp of normalization, indexing, and database design

concepts are all requirements for success in this position.

Database designers emphasize abilities such data modelling, security, and communication to construct the blueprint while focusing on converting business requirements into database schemas. Database developers setup and manage databases, guaranteeing data correctness through SQL queries and optimization, and working together with data analysts and developers to effectively satisfy data needs.

Part 2 Completion: Fill the Blanks Exercises

For each of these exercises, complete each sentence or phrase with the appropriate word or words.

(1) Information is the result of processing raw data to reveal its meaning.

(2) Raw data must be properly formatted for storage, processing and presentation.

(3) Metadata is data about data through which the end - user data are integrated and managed.

(4) A(n) ad hoc query is a spur - of - the - moment question.

(5) Analytical databases focus primarily on storing data used to generate information required to make tactical or strategic decisions.

(6) Data independence exists when it is possible to make changes in the data storage characteristics without affecting an application program's ability to access data.

(7) Data redundancy exists when different and conflicting versions of the same data appear in different places.

(8) A(n) Anomaly develops when all required changes in the redundant data are not made successfully.

(9) Performance tuning relates to activities that make a database operate more efficiently in terms of storage and access speed.

(10) Web and mobile technologies that enable " anywhere, anytime, always on " human interactions are forms of social media.

Part 3 Short Answer Exercises

Consider the data sheet file that follows and then answer the questions pertaining to this table of information. .

ITMD 523	*Advanced Database Management*	Homework 1
Student Name	**Harshal Sawant**	Due Date 09/09/23
Instructor	**Prof. Luke Papademas**	Section 05

PROJECT_CODE	PROJECT_MANAGER	MANAGER_PHONE	MANAGER_ADDRESS	PROJECT_BID_PRICE
21-5Z	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	16833460.00
25-2D	Jane D. Grant	615-898-9909	218 Clark Blvd., Nashville, TN 36362	12500000.00
25-5A	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	32512420.00
25-9T	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	21563234.00
27-4Q	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	10314545.00
29-2D	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	25559999.00
31-7P	William K. Moor	904-445-2719	216 Morton Rd., Stetson, FL 30155	56850000.00

(1) How many records does the file contain? How many fields are there per record?

➔ The file contains 7 records.
Each record has 5 fields: PROJECT_CODE, PROJECT_MANAGER, MANAGER_PHONE, MANAGER_ADDRESS and PROJECT_BID_PRICE.

(2) What problem would you encounter if you wanted to produce a listing by city? How would you solve this problem by altering the file structure?

➔ As the MANAGER_ADDDRESS in the table is combined, it would be difficult to list cities individually. So, to solve this issue we need to sort the address into 4 different fields, such as Street number, City, State and Zip code.

(3) If you wanted to produce a listing of the file contents by last name, area code, city, state or postal zip code, how would you alter the file structure?

➔ We must sort the MANAGER_PHONE column into 2 separate fields such as area code and number. In the same manner, we must separate PROJECT_MANAGER into sub-field of manager's first and lastname. As a result, the necessary contents will be listed in the proper order.

(4) What data redundancies do you detect? How could those redundancies lead to anomalies?

➔ There are 3 different project codes for Holly B. Parker and 2 for George F. Dorts but they both have same respective addresses and phone numbers. This resulted in data redundancy. This data redundancy can cause Insertion, deletion, and update anomalies. For instance, the file has to reflect the change if Holly modifies her address or phone number in any of her three projects. Failure to do so might lead to issues and locating the exact data can be tricky. By reducing duplicate data and enhancing data integrity in databases, normalization is a frequent strategy to address these problems.

(5) Using your <u>own name</u> as the Project Manager, add an additional row to the data table using some fictional but realistic values for the other column entries.

ITMD 523	*Advanced Database Management*	Homework 1
Student Name	**Harshal Sawant**	Due Date 09/09/23
Instructor	**Prof. Luke Papademas**	Section 05

Project_Code	Project_Manager	Manager_Phone	Manager_Address	Project_Bid_Price
29-9D	Harshal D. Sawant	312-885-9643	100 River Rd, Chicago, IL 60616	28228943.00

Part 4 Essays Exercises

Write a brief but complete answer to each of these questions / exercises.

(1) Review the topic of metadata and then discuss what value they provide to the database system.

Metadata refers to the data which describes or give insights about the other data. It offers crucial details on many elements of data, assisting users, programs, and database systems in efficiently managing, comprehending, and using the data.
Metadata serves vital functions in a database system. It provides a description of the data's appearance, assists in keeping it properly organized, and ensures its accuracy. By limiting access, it also makes it easier for users to query the database and protects the data. Even the origin of the data and its evolution through time are tracked via metadata. It acts as a manual for the database, improving its functionality and enabling users to comprehend and use the data safely and easily.

(2) Review and discuss some considerations when designing a database.

Database designing process involves four major stages: Conceptualizing the design, selection of the DBMS, logical mapping of design and physical designing.
Determine the database's purpose and structure, build tables and fields, create primary keys, define data types, use indexes for effective searches, enforce data validation rules, ensure security, plan for scalability, set up backups, design an intuitive user interface, thoroughly test the database, perform routine maintenance, and document the database are all essential steps in database design. Similar to creating a well-organized library or toolkit, or even making a perfect dish like pizza, these factors aid in creating a structured and trustworthy data system.

(3) Discuss the problems associated with file systems. How do they challenge the types of information that can be created from the data as well as the accuracy of the information?

File systems feature problems including poor change detection and organization of data. These issues make it challenging to locate and utilize data properly and might result in errors. This chaos makes it difficult for us to interpret the data and reduces the accuracy of the information. File systems, to put it simply, make it difficult to interact with data and can provide false information.

(4) The Database Management System performs the following functions:

- Data Dictionary Management,
- Data Storage Management,
- Data Transformation and Presentation,
- Security Management,
- Multi - User Access Control,
- Backup and Recovery Management,
- Data Integrity Management,
- Database Access Languages and Application Interface,
- Database Communication Interface.

From the above list, discuss any four functions performed by the DBMS that guarantee the integrity and consistency of the data in the database.

A database management system (DBMS) performs the following four tasks to guarantee the consistency and integrity of data:
1. Data Integrity Management: DBMS prohibits the entry of incorrect or contradictory data into the database by enforcing rules to ensure the accuracy and dependability of data.
2. Security management: This process regulates database access, ensures that only authorized users can make changes, and prevents unauthorized alterations that can jeopardize the consistency of the data.
3. Multi-User Access Control: DBMS controls simultaneous access by several users, avoiding conflicts and guaranteeing that data alterations don't clash with one another, maintaining data consistency.
4. Backup and Recovery Management: DBMS periodically produces backups and aids in data recovery in case of mistakes or failures, ensuring that data is consistent and undamaged even in the face of unanticipated situations.

(5) What are the advantages of having the DBMS between the end user's applications and the database?

There are several benefits to placing a database management system (DBMS) in between the end user's applications and the database. In the beginning, it offers a standardized method of gathering and storing data, making information retrieval and management simpler. Second, by limiting who may access and alter the data, it guarantees data security. Third, a DBMS enables concurrent usage of the database by numerous users without causing problems. Fourth, it reduces the risk of data loss by streamlining the backup and recovery of data procedures. Finally, by implementing data constraints and validation criteria, it improves data integrity by assisting in upholding the correctness and consistency of the information recorded in the database. Simply said, a DBMS functions as a useful middleman who simplifies, protects, and organizes our interactions with our data.

ITMD 523	*Advanced Database Management*	Homework 1
Student Name	**Harshal Sawant**	Due Date 09/09/23
Instructor	**Prof. Luke Papademas**	Section 05