

Page No.	
Date	

College Name :

DELHI GLOBAL INSTITUTE OF TECHNOLOGY

Name : BAZGHA RAZI

Course Code : PCC-CSE-201 G

Subject : Database Management System

Session : 2019 - 2023

Ans 16) Of data warehousing is process for collecting and managing data from varied sources to provide meaningful business insights.

A data warehouse is typically used to connect and analyze business data from heterogeneous sources.

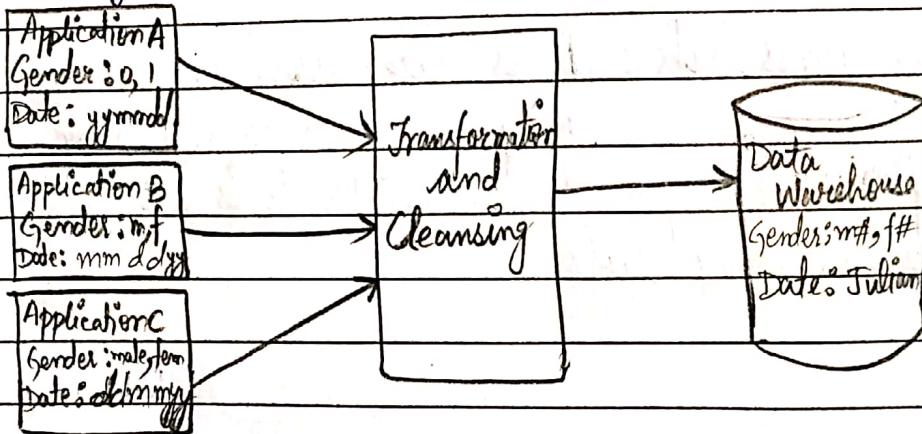
It is electronic storage of a large amount of information by a business which is designed for query and analysis instead of transaction processing.

It is a process of transforming data into information and making it available to users in a timely manner to make a difference.

It is an architectural construct of an information system which provides users with current and historical decision support information which is difficult to access or present in the traditional operational data store.

A data warehouse provides a new design which can help to reduce the response time and helps to enhance the performance of queries for reports and analytics.

Data warehouse benefits users to understand and enhance their organisation's performance.



Components of data warehouse are as follows:

- i) Load manager: It is also called the front component. It performs with all the operations associated with the extraction and load of data into the warehouse. These operations include transformations to prepare the data for entering into the datawarehouse.
- ii) Warehouse manager: It performs operations associated with the management of the data in the warehouse. It performs operations like analysis of data to ensure consistency, creation of indexes and views, generation of denormalization and aggregations, transforming & merging of source data and archiving and backing-up data.
- iii) Query Manager: Query manager is also known as backend component. It performs all the operations related to the management of user queries. The operations of this data warehouse components are direct queries to the appropriate tables for scheduling the execution of queries.
- iv) End-user access tools: It is categorised into five different group i.e., data reporting, query tools, application development tools, EIS tools, OLAP tools and data mining tools.

Data warehouse used for:

- i) Airline: In this, it is used for operation purpose like crew assignment, analyses of route profitability, frequent flyer program promotions, etc.
- ii) Public sector: In this sector, it is used for intelligence gathering. It helps government agencies to maintain and analyze tax records, health policy records, for every individual.
- iii) Retail chain: In this, it is used for distribution and marketing. It also track market customer buying pattern, promotions and also used for determining pricing policy.
- iv) Telecommunication: It is used in this sector for product promotions, sales decisions and to make distribution decisions.
- v) Insurance and investment sector: In this sector, it is used primarily to analyze data patterns, customer trends, and to track market movements.

Ans 1) A web database is a wide term for managing data online. A web database gives you the ability to build your own database / data storage without you being a database guru or even a technical person.

Examples → The web is a distributed information system base on hypertext.

→ Most web documents are hypertext documents formatted via HTML.

→ HTML documents contain text along with font specification, which can be associated with region of the text

→ Bank, airline, university course registration, etc.

Applications of Web database

The manner in which many organizations use a database in conjunction with a website is a good example.

It helps us to serve up banner advertisements on web pages. We don't like them any better than you do, but the fact remains that they are a popular application for web databases, which can be used to store advertisements, and retrieve them for display by a web server.

It helps you to make the website more useful in ways not even related to the membership list. You may publish an electronic newsletter that has a children's section in each issue containing a relevant quiz.

Advantages of Webdatabase

Flexible use : Using webdatabase is very beneficial because it allows your business to be flexible.

You only pay for the amount of storage that you use. You need not worry about purchasing servers as you go or eliminating them when they are no longer needed.

Save Money : One of the advantages of webdatabase is that it can save your business money.

When you don't need to buy a software program for your business, this could result in a major savings overall.

Technical Support : By using Webdatabase you can shift the burden of technical support to someone else. Paying a company for access to an online web database include Technical support. If the database has problems, you simply contact the company and the staff handles it.

Access : Having access to the database at all times from multiple locations is another major advantage of webdatabase. With an online database, you could theoretically access the information in the database from any computer.

Disadvantages of Web database

No internet connection : Online means... online. If you don't have a reliable internet connection to support your live online database software, you don't have a reliable access to your data. In simple words, we can say that no internet connection means no access.

Per user pricing : In terms of pricing, an installed solution represents one big lump sum when you acquire the software. You have to be careful when you price your online database solution since the price usually relates to the number of users and applications you will require. The bigger the team, the pricier it gets. Same for your apps, the more you need, the more you pay.

Cost of extra storage : With virtually all vendors, you will face a cap when it comes to storage space and number of items you can create. If you are only creating text items without large attachments, don't lose sleep over this, you will never see the end of the storage your vendor is allowing you.

Ans 2 b) Basis of Comparison

Indexing

Hashing

Definition

Indexing is a data structure technique to efficiently retrieve records from the database files based on some attributes on which the indexing has been done.

Hashing is an effective technique to calculate the direct location of a data record on the disk without using index structure.

Functionality

Indexing uses data reference that holds the address of the disk block with the value corresponding to the key.

Hashing uses mathematical functions called the hash functions to calculate direct locations of data records on the disk.

Working

It does not work for large databases.

It works well for large databases.

Uses

Indexing is used to optimize the performance of a database by minimizing the number of disk accesses required when a query is processed.

Hashing is used to index and retrieve items in a database as it is faster to search that specific item using the shorter hashed key instead of using its original value.

Sorting of address

Addresses in the memory are sorted according to a key value called the primary key.

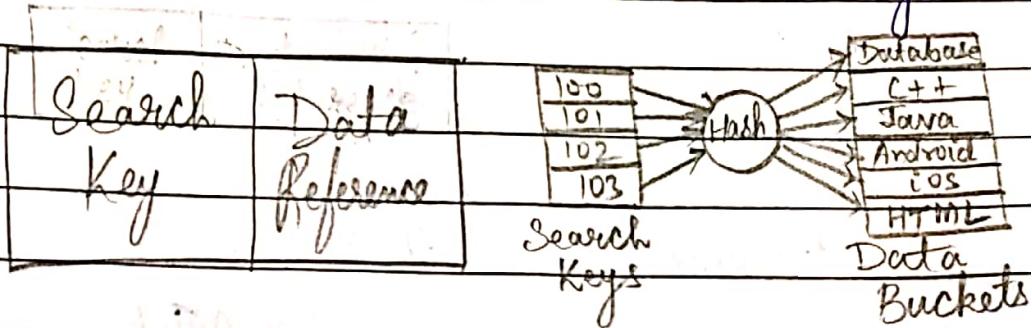
Addresses are always generated using a hash function on the key value.

Performance

It stores the data in a sorted form when there is any operation performed which decreases its performance. Performance of hashing will be best when there is a constant addition and deletion of data.

Example

It can be created using some database columns. It uses buckets and search keys



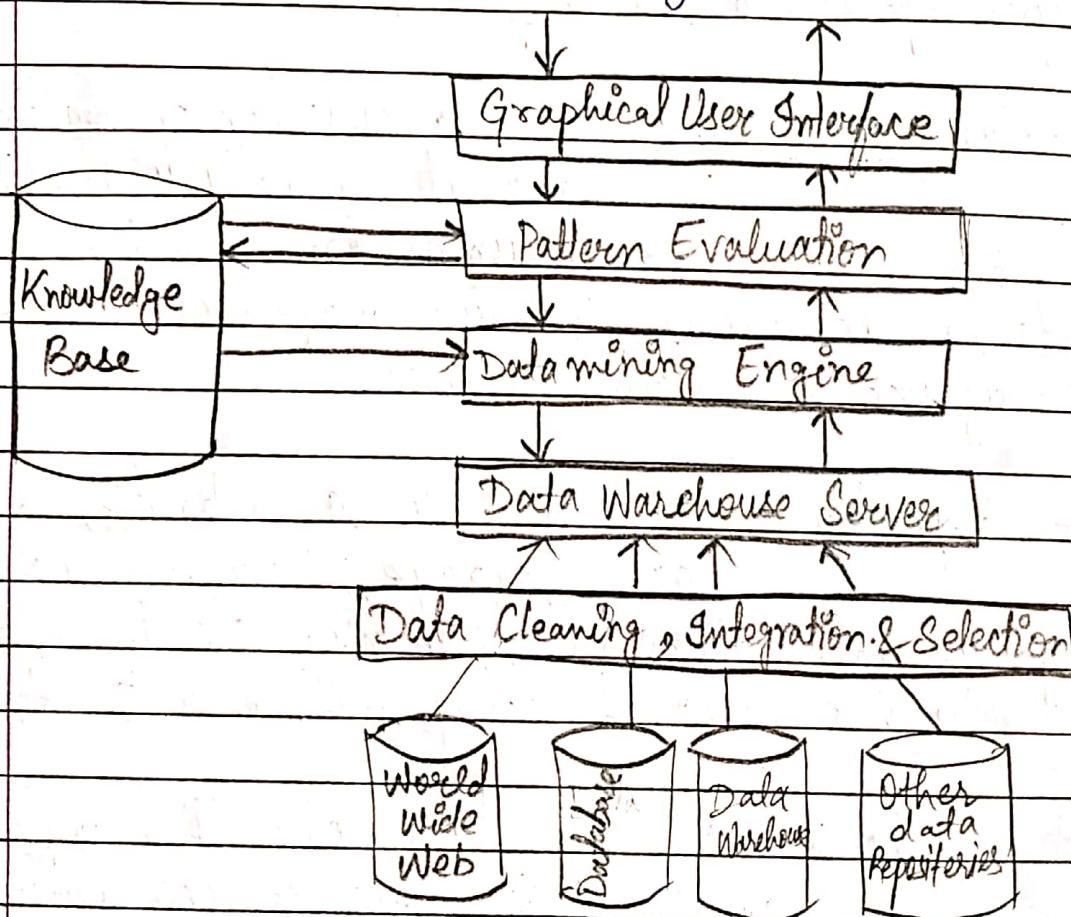
Ans 3(b) Data Mining : It is the process of extracting information to identify patterns, trends and useful data that would allow the business to take the data - driven decision from huge sets of data. It is called data mining. It is the process used by organisations to extract specific data from huge databases to solve business problems. It primarily turns raw data into useful information.

Types of Data Mining

- i) Relational Database : It is a collection of multiple data sets formally organised by tables, records and columns from which data can be accessed in various ways without having to recognize the database tables.
- ii) Data Repositories : It refers to a destination for data storage. However, many IT professionals utilize the term more clearly to refer to a specific kind of setup within an IT structure.
- iii) Data Warehouse : It is the technology that collects data from various sources within the organisation to provide meaningful business insights. The huge amount of data comes from multiple places such as Marketing and finance.

- iv) ~~Relational~~ Object-~~Relational~~ Database : A combination of an object-oriented database model and relational database model is called Object-~~Relational~~ model.
- v) Transactional Database : It refers to a database management system that has the potential to make a database transaction if it is not performed appropriately.

Architecture of Data mining

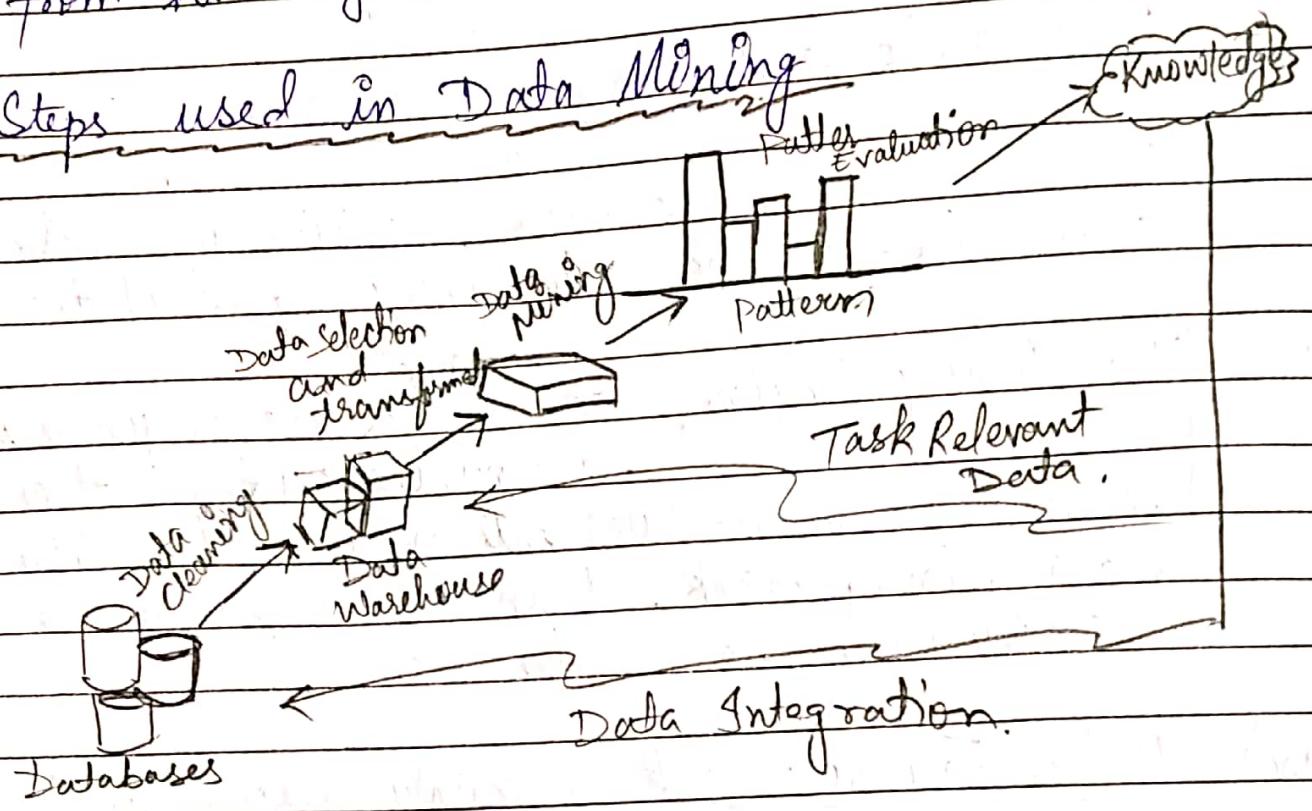


Data mining's architecture contains too many components. That is data source, data warehouse server, data mining engine and knowledge base.

- i) Data Source : There are so many documents present. Such as database, datarehouse, world wide web (www). These are the actual source of data. Data may reside even in plain text files or spreadsheets. World wide Web is another big source of data.
- ii) Datawarehouse Server : It contains the actual data that is ready to be processed. Hence, the server handles retrieving the relevant data. It is based on the data mining request of the user.
- iii) DataMining Engine : It is the core component in data mining system. It consists of a number of modules that we used to perform data mining task. It includes, association, classification, prediction, etc.
- iv) Pattern Evaluation : It is mainly responsible for the measure of the interestingness of the pattern. We use threshold value for this. It interacts with data mining engine. Main work of this is to search towards the interesting patterns.
- v) Graphical User Interface : It is used to communicate between the user and the data mining system. It helps the user to use the system easily & efficiently. It displays the result in an easily understandable manner.

v) Knowledge Base: It is the most beneficial part of data mining. It guides the search for the result patterns. It contains data from user experiences i.e., useful in the process of data mining. Data mining engine might get inputs from knowledge base.

Steps used in Data Mining



Step 1) Data Cleaning: The noisy and irrelevant data is removed from the collection.

Step 2) Data Integration: The heterogeneous data from multiple source is combined in a single source

Step 3) Data Selection: In this step, ~~not~~ relevant data is decided and retrieved from the data collection.

Step 4) Data Transformation: Data is transformed into appropriate form i.e required by the data mining procedure

Step 5) Data Mining: In this, some useful or intelligent techniques are applied to extract meaningful information.

Step 6) Pattern Evaluation: Increasing patterns that represent knowledge based on the given measures.

Step 7) Knowledge Representation: It is defined as technique which utilizes visualization tools to represent data mining results.

Uses of Data Mining

i) Used in business: Data mining help a company to decrease its costs, increase revenues or derive insights from the behaviour and practices of its customers.

ii) Used in finance: Data mining techniques allows users to determine and assess the factors that influence the price fluctuations of financial securities.

iii) Fraud

Page No.	
Date	

- iii) Fraud Detection : It helps to include the collection of sample records on the basis of fraudulent or non-fraudulent. And after collecting it checks whether the data is fraudulent or not.
- iv) Compress data in valuable form or information
- v) Market analysis : This technique will allow the retailer to understand the purchase behaviour of buyers which will help them in changing the stores setup according to their sale.