29 Oct 2024

DATABASE FUNDAMENTALS

Databases:
A database is a shared collection of logically related data and description of these data, designed to meet the information needs of an organisation.

Data storage: A database is used to store large amount of structured data, making it easily accessible, searchable and retrievable.

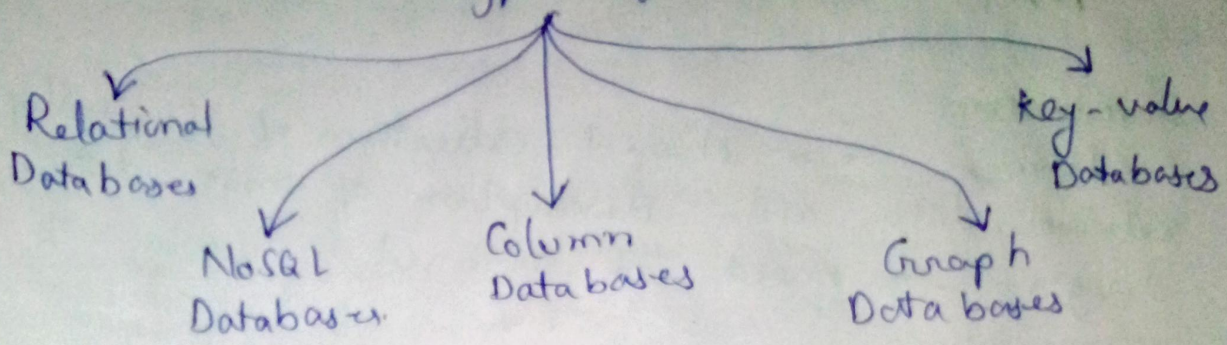
Data Analysis: A database can be used to perform complex data analysis, generate reports and provide insights into the data.

Record keeping: A database is often used to keep track of important records, such as financial transactions, customer information and inventory levels.

Web Applications: Databases are an essential components of many web applications providing dynamic content and user management.

properties of an ideal Database:
1. Integrity: accuracy + consistency
2. Availability
3. Security
4. Independent of application.
5. Concurrency

Types of Databases



1. Relational Databases:
Also known as SQL databases, these databases use a relational model to organize data into tables with rows and columns. example. MySQL, Postgre, Oracle, SQLServer, Microsoft Access.

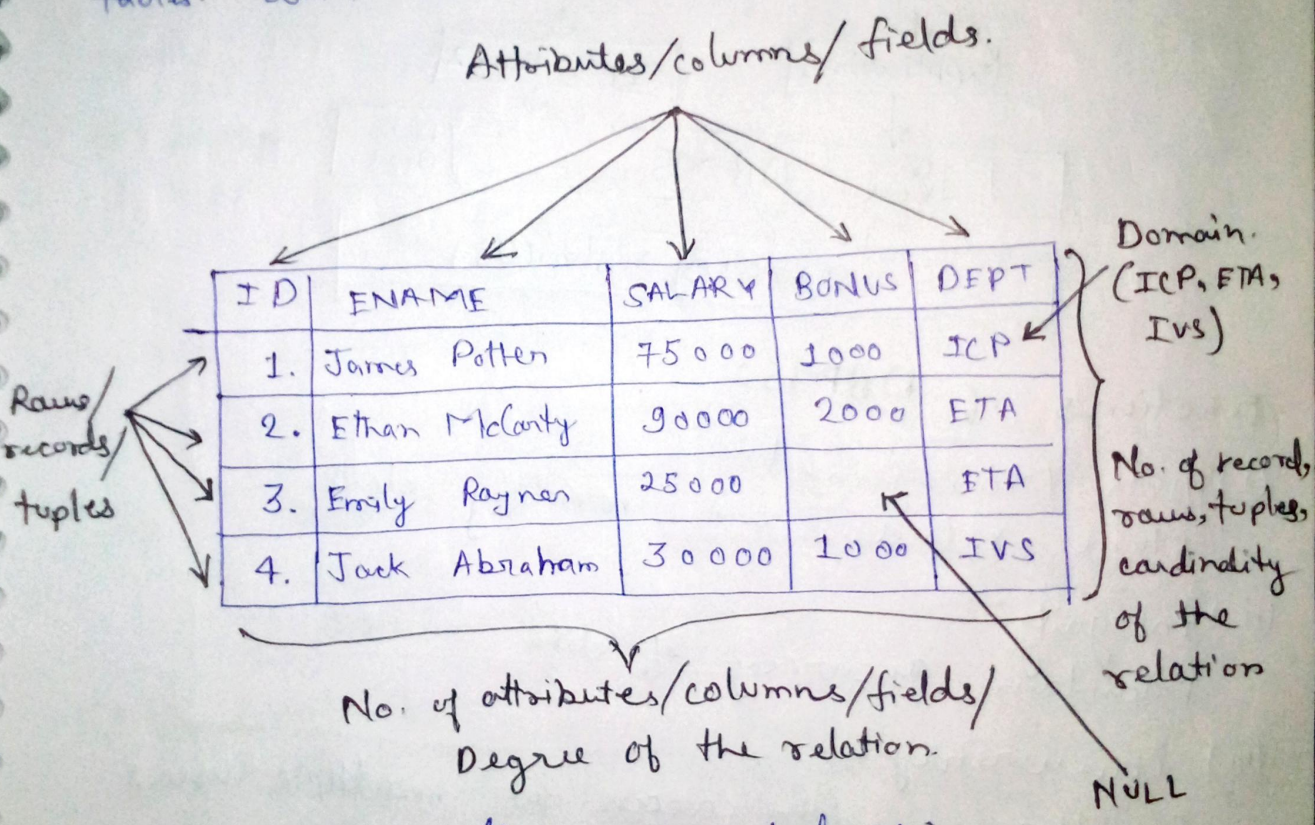
2. NOSQL Databases:
These databases are designed to handle large amount of unstructured or semi-structured data, such as documents, images or videos example MongoDB.

3. Column Databases:
These databases store data in columns rather than rows, making them well suited for data warehousing and analytical applications. example. Amazon Redshift, Google BigQuery.

4. Graph Databases:
These databases are used to store and query graph-structured data, such as social network connections or recommendation system example. (Neo4j, Amazon Neptune).

5. Key-Value Databases:
These databases store data as a collection of keys and values making them well-suited for caching and simple data storage needs example. Redis and Amazon DynamoDB

Relational Databases:
Also known as SQL databases, these databases use a relational model to organize data into tables with rows and columns.

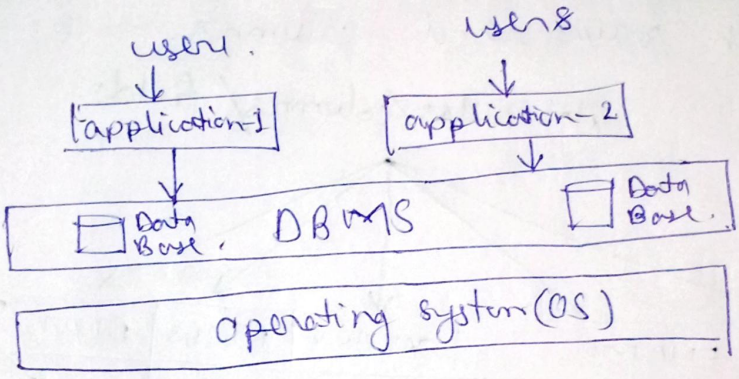


ID	ENAME	SALARY	BONUS	DEPT
1.	James Potter	75000	1000	ICP
2.	Ethan McCarty	90000	2000	ETA
3.	Emily Rayner	25000		ETA
4.	Jack Abraham	30000	1000	IVS

Relation is usually represented as:
Employee (ID, ENAME, SALARY, BONUS, DEPT)

Data Base Management System (DBMS):

A database management system (DBMS) is a software system that provides the interfaces and tools needed to store, organize and manage data in a database. A DBMS acts as an intermediary between the data base and the application, or users that access the data stored in the database.



Functions of DBMS:

(i) Data Management:
store, retrieve and modify data.

(ii) Integrity:
Maintain accuracy of data

(iii) Concurrency:
Simultaneous data access for multiple users.

(iv) Transaction:
Modification to database must either be successful or must not happen at all

(v) Security:
Access to authorized users only.

(vi) Utilities:
Data import/export, user management, backup, logging.

Database keys:
A key in a database is an attribute or set of attributes that uniquely identifies a tuple (raws) in a table. keys play a crucial role in ensuring the integrity and reliability of a database by enforcing unique constraints on the data and establishing relationships between tables.

(i) Super key:
A super key is a combination of columns that uniquely identifies any raw within a relational database management system (RDBMS) table. example. Roll no., email, roll no+ name

Roll no.	Name	Branch	email
[redacted] 1	John	CSE	John@gmail.com
[redacted] 2	Ram	EEE	ram@gmail.com
[redacted] 3	Adil	ME	adil@gmail.com

(ii) Candidate key:
A candidate key is a minimal super key, meaning it has no redundant attribute. In other words, it's the smallest set of attributes that can be used to uniquely identify a tuple (raw) in a table. example. Roll no, email.

(iii) Primary key:
A primary key is a unique identifier for each tuple in a table. there can only be one primary key in a table, and it cannot contain null values.

→ Null
→ repite - duplicate
→ numerical
→ Small
→ constant

(iv) Alternate key:
An alternate key is a candidate key that is not used as the primary key.

candidate key - Primary key = alternate key

(v) Composite key:
A composite key is a primary key that is made up of two or more attributes. Composite keys are used when a single attribute is not sufficient to uniquely identify a tuple in a table.

student_table

sid	name	email	phone

course_table

cid	Name	price	instructor

enrollment table.

sid	cid	date	payment

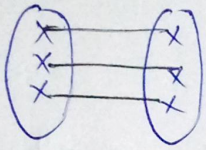
Here sid + cid is composite key

(vi) Surrogate key:
adding the new column in the table when no one column or combination of column are able to become a primary key.

(vii) Foreign key:
A foreign key is an primary key from one table that is used to establish a relationship with another table. example sid and cid

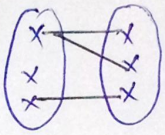
Cardinality of Relationships:

Cardinality in database relationships refers to the number of occurrences of an entity in a relationship with another entity. Cardinality defines the number of instances of one entity that can be associated with a single instance of the related entity.



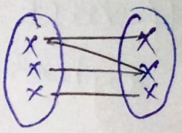
1:1
One-to-one relationship

e.g:
person → driving license
(one table require)



1:N
One-to-many relationship

e.g:
student → college branch
(two tables required)



M:N
Many-to-many relationship

e.g:
student → courses
(three table requir)

Drawbacks of Databases:

(i) Complexity:
setting up and maintaining a database can be complex and time consuming, especially for large and complex systems.

(ii) Cost:
The cost of setting up and maintaing a database, including hardware, software and personnel can be high.

(iii) Scalability:
As the amount of data stored in a database grows, it can become more difficult to manage, leading to performance and scalability issues.

(iv) Data Integrity:
Ensuring the accuracy and consistency of data stored in a database can be challenge, especially when multiple users are updating the data simultaneously.

(v) Security:
Securing a database from unauthorized access, and protecting sensitive information can be difficult, especially with the increasing threat of cyber attacks.

(vi) Data Migration:
Moving data from one database to another or upgrading to a new database can be a complex and time-consuming process.

(vii) Flexibility:
The structure of a database is often rigid and inflexible, making it difficult to adapt to changging requirements or to accommodate new types of data.