Introduction to DBMS & Relational Model

TABLE OF CONTENTS

* 40 days hard challenge :

1st Hard day challenge :

1. Assignment (MCQs)
2. Do a quick revision

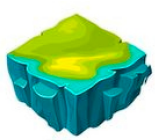
DSA

Not a DSA fan?

SQL (NON-DSA)

Dive into SQL's path to success

- **How should you tackle SQL module?**

	Task-1	Task-2	Task-3	Task-4	
					
YOU	Attend Lectures	Increase psp	Contest	Mock	Mission Accomplished

>90 %

Watch Recording

X 2

> 90 %

20 Questions

10 MCQs

10 Query

Curriculum..

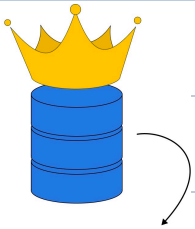
1. Intro to DBMS and SQL
2. Keys
3. Crud x 2
4. Joins x 2
5. Aggregates
6. Subqueries
7. Indexes
8. Transactions x 2
9. Schema Design x 2

What isn't covered

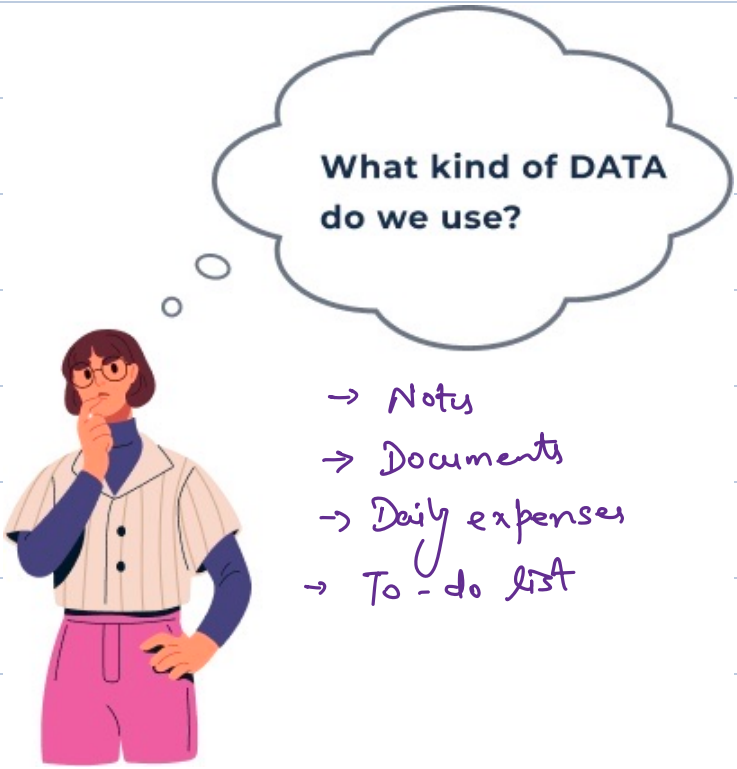
1. Distributed databases
2. Scalability and related concepts like NoSQL
3. Database Sharding
4. Master Slave architecture
5. Database Replication

→ HLD

Why do we talk so much about DATA always?



DATA is king



Applications :

To-Do

Excel

students.csv

H18 | fx

	A	B	C	D
1	**ID**	**NAME**	**PSP**	**ATTENDANCE**
2	1	Himanshu	80	85
3	2	Rahul	75	90
4	3	Krish	95	95
5	4	Rahul	92	85
6	5	Rohit	80	88

Question

Find average psp of student corresponding to their batches.

→ Open / Read

→ Iterate

→ Calculation

TC = O(N)

Drawbacks :

1. Inefficient

2. Data Integrity

students.csv

K16 fx

	A	B	C	D
1	ID	NAME	PSP	ATTENDANCE
2	1	Himanshu	80	85
3	2	Rahul	75	90
4	3	Krish	95	95
5	4	Rahul	Topper	85
6	5	Rohit	80	88

→ Data Inaccuracy

→ In case of PSP clm we expected numerical data but got str as well.

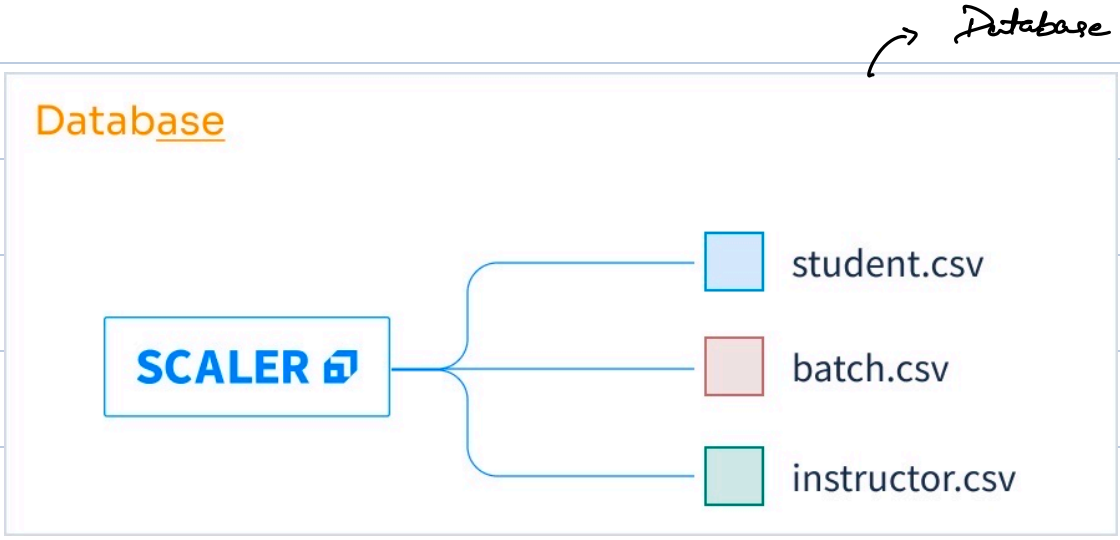
3. Concurrency

⇒ If more than 1 person working on same file at same time.

4. Security Issues

→ Data Security

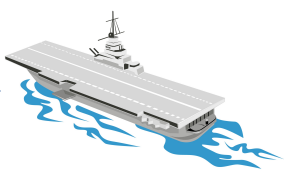
What is Database?



Army Base

Air Base

Naval Base



Management

Database Management System (DBMS)

- A DBMS as the name suggests is a software system that allows to efficiently manage a database.

- A DBMS allows us to create the following :

  Create

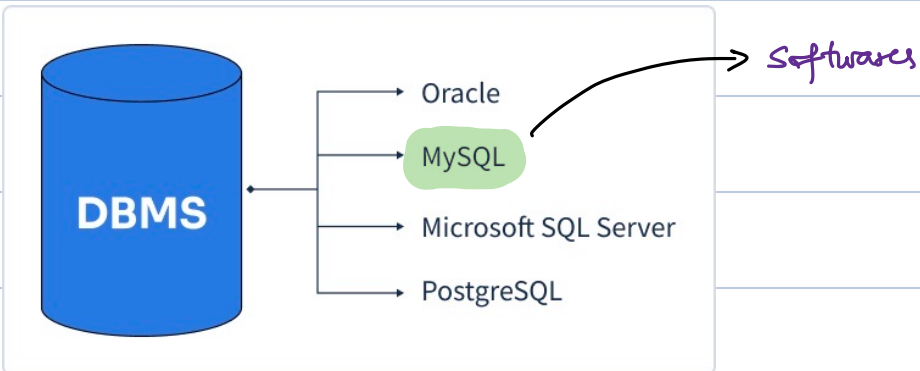
  Read

  Update

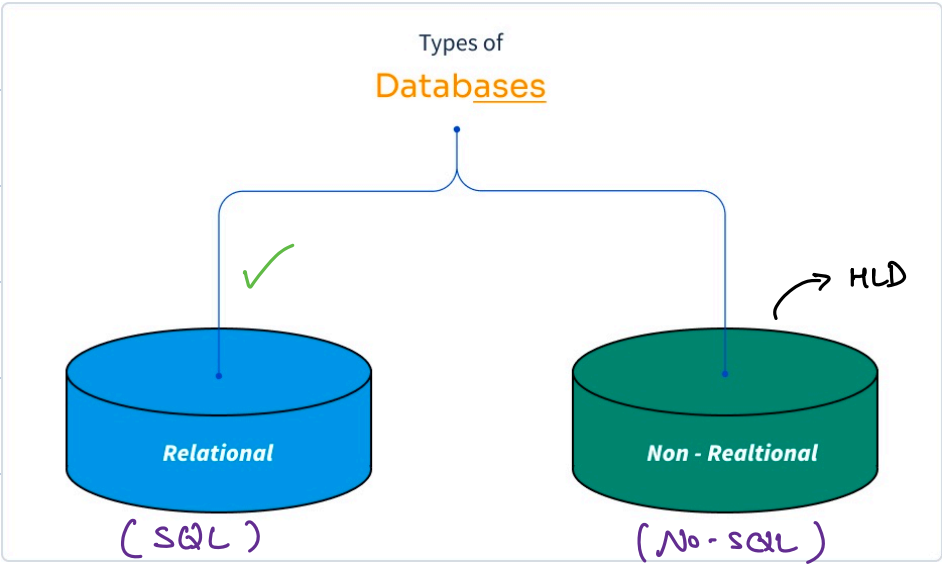
  Delete

- It allows us to define rules to ensure :

  1. Data Integrity
  2. Security and
  3. Concurrency



Types of Databases



1. Relational

The data is being stored in form of inter-related tables

Related

Students

id	name	psp	attendance	b_id
1	Himanshu	80	85	2
2	Rahul	75	90	1
3	Krish	95	95	1
4	Rahul	92	85	2
5	Rohit	80	88	1

Batch

id	name
1	A
2	B
3	C
4	D
5	E

2. Non - Relational

- Don't store data in form of tables.
- Store data in form of documents, key-value pairs, graphs, etc.
- We will talk more about them in the HLD Module.

Properties of RDBMS

1. Relational Databases represent database as a collection of tables with each table storing information about something.

Students Batches Instructors

2. Every row is unique.

Students

name	psp	attendance	b_id
Himanshu	80	85	2
Rahul	92	85	2
Krish	95	95	1
Rahul	92	85	2
Rohit	80	88	1

Question

Find psp of Rahul.

Duplicacy leads to ambiguity.

Students

id	name	psp	attendance	b_id
1	Himanshu	80	85	2
2	Rahul	92	85	2
3	Krish	95	95	1
4	Rahul	92	85	2
5	Rohit	80	88	1

3. A column should have all values of same data type.

Students

int str int

id	name	psp	attendance	b_id
1	Himanshu	80	85	2
2	Rahul	75	90	1
3	Krish	95	95	1
4	Rahul	Topper	85	2
5	Rohit	80	88	1

RDBMS can help us hardbound columns to store data of a single data type.

4. All values / cell should be atomic.

Students

id	name	psp	phone.no	b_id
1	Himanshu	80	956453789	2
2	Rahul	75	906453875	1
3	Krish	95	829376769, 806122348	1
4	Rahul	92	806122348	2
5	Rohit	80	762766434	1

5. The sequence of column is not guaranteed by RDBMS.

Note : MySQL preserves the order.

Students

id	Name	psp	attendance
1	Rahul	82	85

6. The sequence of row is not guaranteed.

Note : MySQL preserves the order.

7. Name of every column has to be unique.

Students

id	name	psp	phone.no1	phone.no2
1	Himanshu	80	829376769	72
2	Rahul	75	956453789	Null
3	Krish	95	906453875	Null
4	Rahul	92	806122348	98
5	Rohit	80	762766434	78

Question

id = 4

What is the phone number of Rahul?

→ To avoid ambiguity we have unique names to each columns.

Keys

All rows are not unique

Students

name	psp	attendance	b_id
Himanshu	80	85	2
Rahul	92	85	2
Krish	95	95	1
Rahul	92	85	2
Rohit	80	88	1

Students

id	name	psp	attendance	b_id
1	Himanshu	80	85	2
2	Rahul	92	85	2
3	Krish	95	95	1
4	Rahul	92	85	2
5	Rohit	80	88	1

Question

Update psp of Rahul to 100.

Definition : **Keys helps us to uniquely identify a row.**

Types of Keys

1. Super Keys
2. Candidate Keys
3. Primary Keys
4. Composite Keys
5. Foreign Keys

1. Super Keys

Students

id	name	psp	attendance	b_id
1	Himanshu	80	85	2
2	Rahul	75	90	1
3	Krish	95	95	1
4	Rahul	92	85	2
5	Rohit	80	88	1

1. Can ' name ' column uniquely identify row?
2. Can ' batch ' column uniquely identify row?

Column Name	Super Key
name	✗
psp	✗
attendance	✗
psp, name	✗
id	✓
id, name, psp, b_id	✓
b_id	✗

Definition : **It uses a column / combination of columns to uniquely identify a row.**

- *In case of super key it can use redundant columns.*

Quizzes :

1. a) {Student_id, Course_id} ✓
   b) {first_name, last_name} ✗
   c) {Age, Course_name} ✗
   d) {last_name, Course_id} ✗

2. a) {Student_id, Course_id} ✓
   b) {first_name, Age} ✗
   c) {last_name, Age} ✗
   d) {Course_id, Course_name} ✗

3. a) {Student_id, first_name} ✓
   b) {Student_id, Age} ✓
   c) {Student_id, last_name} ✓
   d) ✓ All of the above

Extra resources : Recordings | Typed Notes

1. String
2. Number
3. Date
4. MySQL Data Types
5. Window Functions
6. Query Executation and Optimization
7. IF, IFNULL, CASE, COALESCE
8. ER Diagrams
9. Database Normalization Forms

Announcement

- **What *NEXT?***

1. Assignment / Homework — **MCQ** → 10 - 15 mins
2. Read Notes (Data Types)