SQL

Relevel
by Unacademy

Problems

- Data is huge.
- Management
- Slow processing
- Data security
- Difficult to input & fetch the data out
- Expansion

Problems

→ The traditional data storage are having their own limitations regarding storage and hence amazon would need more and more of such storage options to preserve their data.

→ Logical arrangment & collection of the data is not possible. They would face issue with data ingestion and retrival

*→ Same view : Mismatch in analysis and the project outcome

Problem

(Preserve data) → Traditional Data storage

→ Excel files
→ PDF files
→ Pen drives
→ Hard paper documents

AMAZON

(details of delivery patner ✔)

[Delivery partner]

USA

INDIA

E-comm

C1

C2

(Laila)

(majnu)

one-project

[Product]

[vendor]

[vendor specific information]

orders

[customer]

[Email, password, Payment method, last order, address]

→ Time loss — Input, retrive, search

RDBMS : Relational Database Management System

Software Infrastructure

- Server (central storage)
- Client (GUI) (Software that's running on user's laptop to connect to the server)
  - graphical user interface

Photographs
- parents
- college days
- Bewaffa g.f
- myself

(Shreya) {username, password}

INDIA

(Sumit)

username, password

client (software)

SQL

Server
Data Center
(Data warehouse)

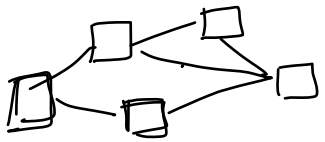
(folder)
- E-comm
- GO
- ENT
- WEB

files { order, customer, vendor, product, delivery partner }

Data Engineers : Core takers of the data center

SQL : Communication Language which act as bridge between the user and the server.

folder - file (Database) (tables)



Snow flake schema

Company Database ←

multiple files (tables)

Employee table → fields

E.id	Name	Joining Date	Sal	Dep. Id

Records (rows)

forigen field

(Dep.id)

Department table

Dep.id	Dep Name	# of employees
	HR	10

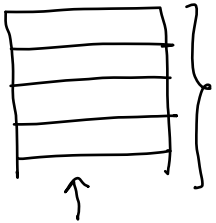
primary field (unique value no nulls)

(Snow flake schema)

(forigen field + primary field)

Relationship

- a particular primary field is known as forigen field if found in any other table.

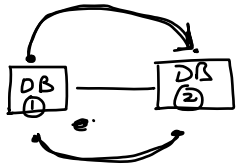


(SQL) → TCL (Transaction control Lang) { commit → , rollback → }

DDL (Data definition Lang)

DML (Data manipulation Lang)

DCL (Data control Lang)



English { ≡ } segregated based on the purpose of usage

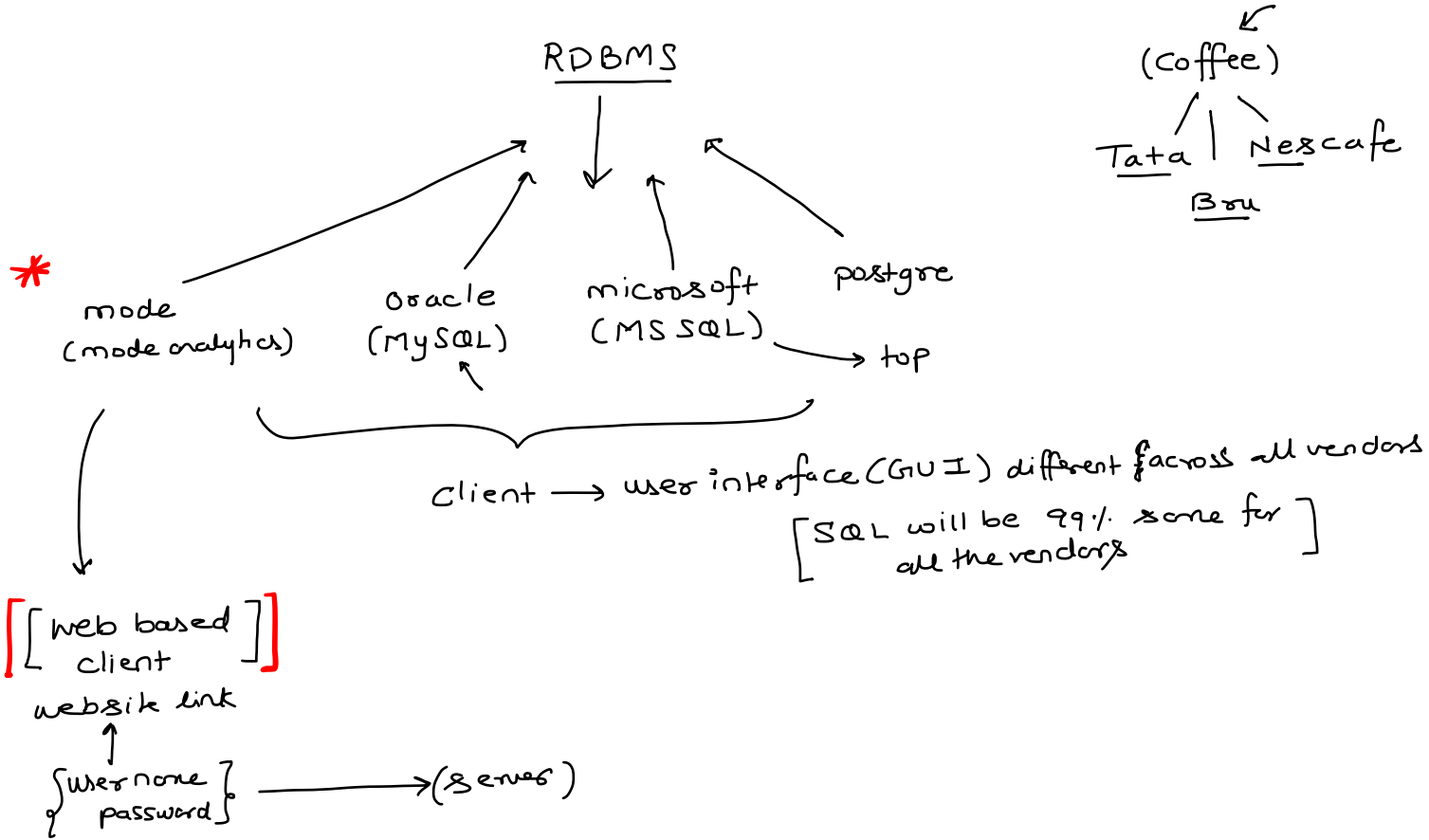
DDL : Data Definition Lang : Data Engineering team to create, delete, replace a database / table in the data center (server) { Create, Alter, drop, truncate, rename. }

DE

DCL : Data control Lang : These are the set of commands used by Data Engineering team to provide the access ~~to~~ or to revoke the access of a user to the server { grant, revoke }

DML : Data manipulation Lang : These are used by BA/DA to manipulate the data available in the server to find the answers of various questions. { Select, from, insert, update }

BA/DA



(SQL is a case insensitive Language)

Data warehouse

↓

(Multiple Database)

(DB-1) (DB-2) (DB-3)

employee
department
project

① Select a database out of all databases available in the datawarehouse, in which we want to write the SQL query and find out various answers.

E.id	Name	age	Emp	Depid

Use DB-2 ; → we use this to let the editor understand that we are done writing a query (end of the query)

② Now we can query any table from this database

Select * → all fields/records

from employee ; → following table.

I want to Select

```
Select *
from sessionwithSumit.employee-data;
```

Name of the database

table name