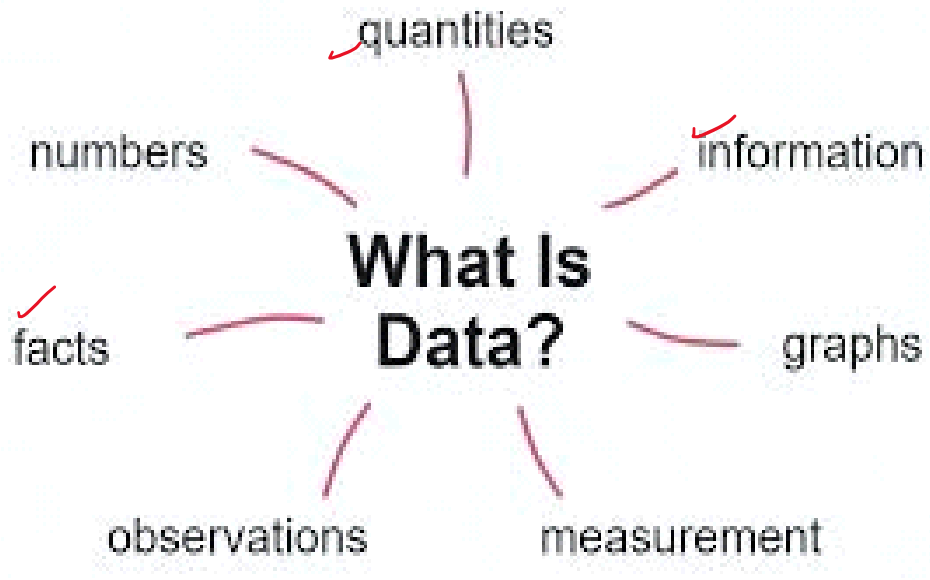
What is Data?

Saturday, July 19, 2025 2:34 PM



- Everything around us is data:
  numbers, text, images, videos.
- Data alone is just raw facts -
  we organize it to get information.

Where do you store the data?

- Based on the format of data you store
- Tabular ⇒ Excel
- text ⇒ document
- image ⇒ files etc

Database

Organised collection of Data stored in an electronic format

→ Store millions of orders safely
→ If many people are working at the same time. Can quickly get you the data.

Day 1 10 orders friday.

Name
Add
Cookie Type
qty
Price
feedback
phone no
delivery

↓ viral!

100 orders friday.

↓

5000 - 15000

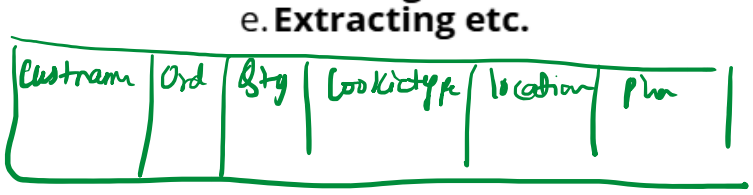
Online.

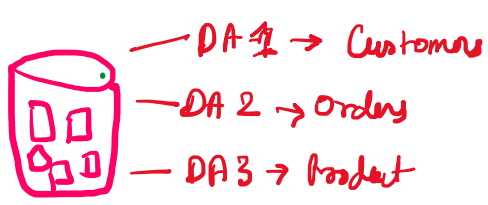
Problems with Excel sheet

- file will be slow to open
- You need complex formulas to do the analysis
- 2 people open the file & do work there is no history / log & not secured.
- One small mistake - can break whole file.

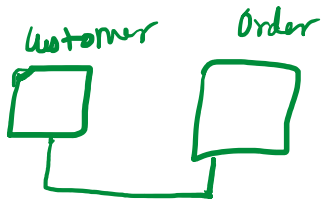
What is SQL?

- SQL (pronounced "ess-que-el") stands for Structured Query Language
- SQL is used to communicate with a database.
- It is the standard language for relational database management systems
- The standard SQL commands such as "Select", "Insert", "Update", "Delete", "Create", and "Drop" can be used to accomplish almost everything that one needs to do with a database.
- SQL programming can be used to perform multiple actions on data such as :
  a. Querying
  b. Inserting
  c. Updating
  d. Deleting
  e. Extracting etc.

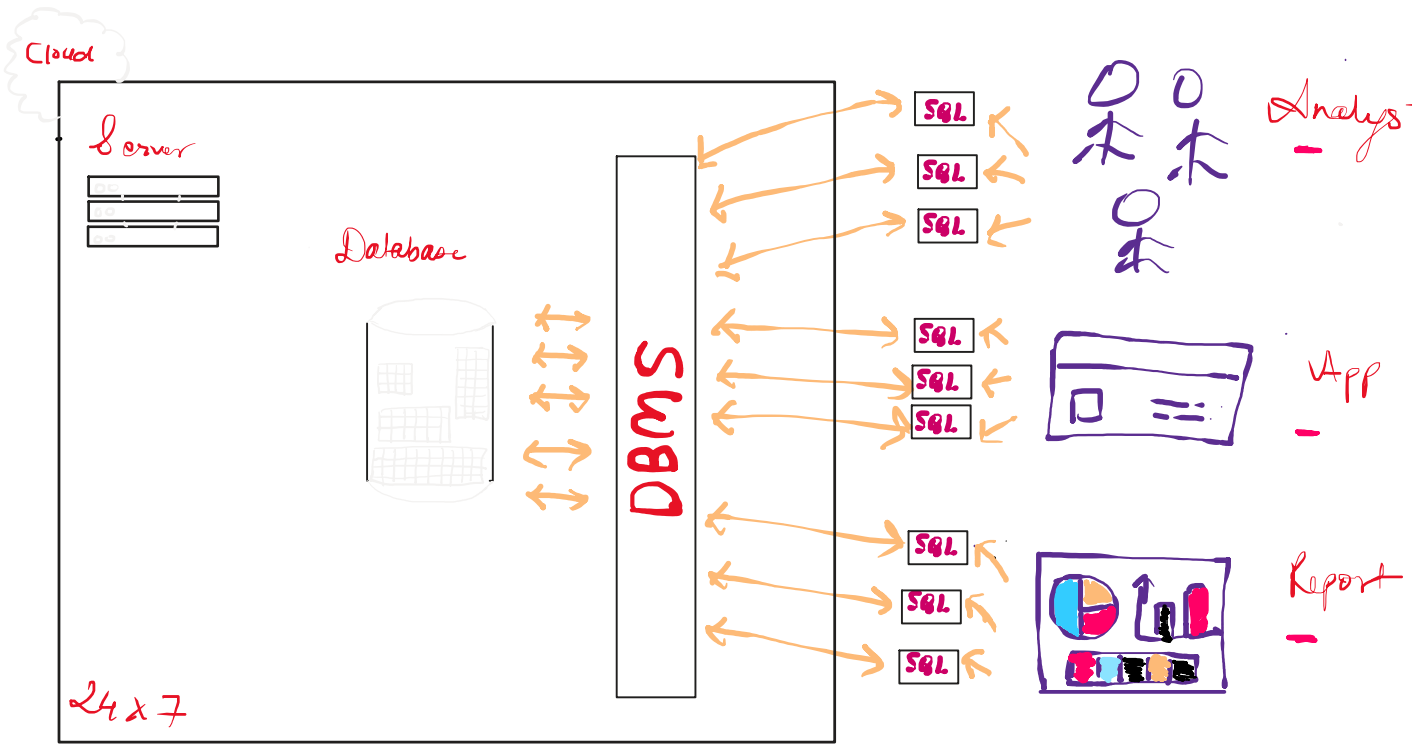


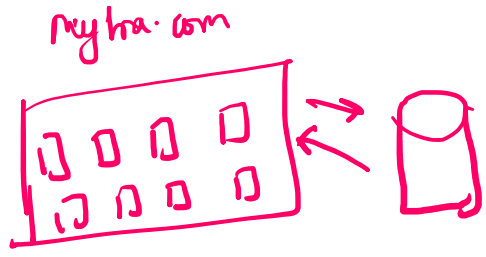


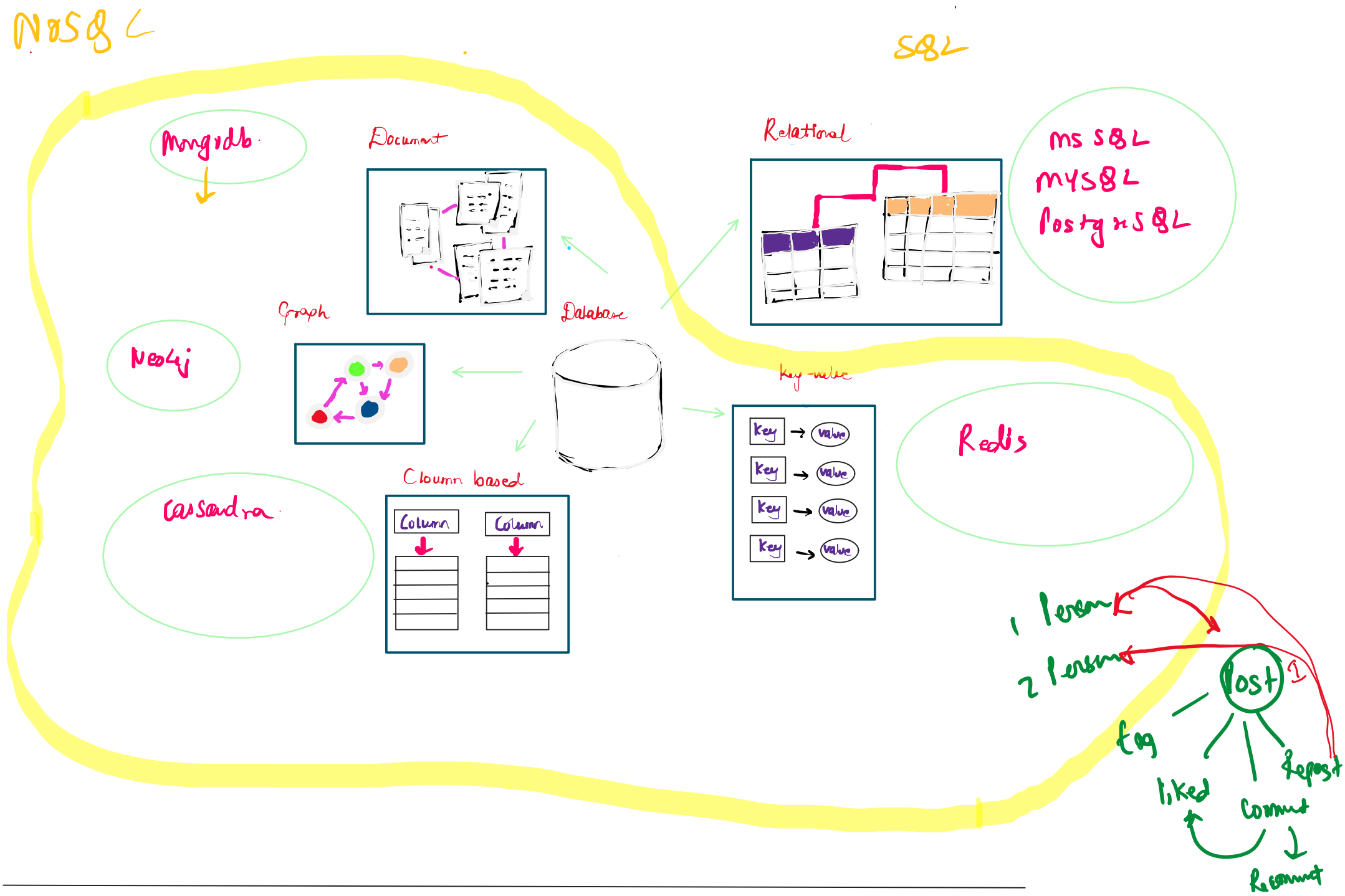
Customer behaviour



Excel	Database
Good for small data (up to a million rows)	Handles billions of rows
Good for analysis, charts, reports	Good for storing, querying, relationships
One person edits at a time	Many users can edit together
Files stored locally	Databases run on servers, accessible anywhere
Not ideal for complex relationships	Designed for multiple related tables

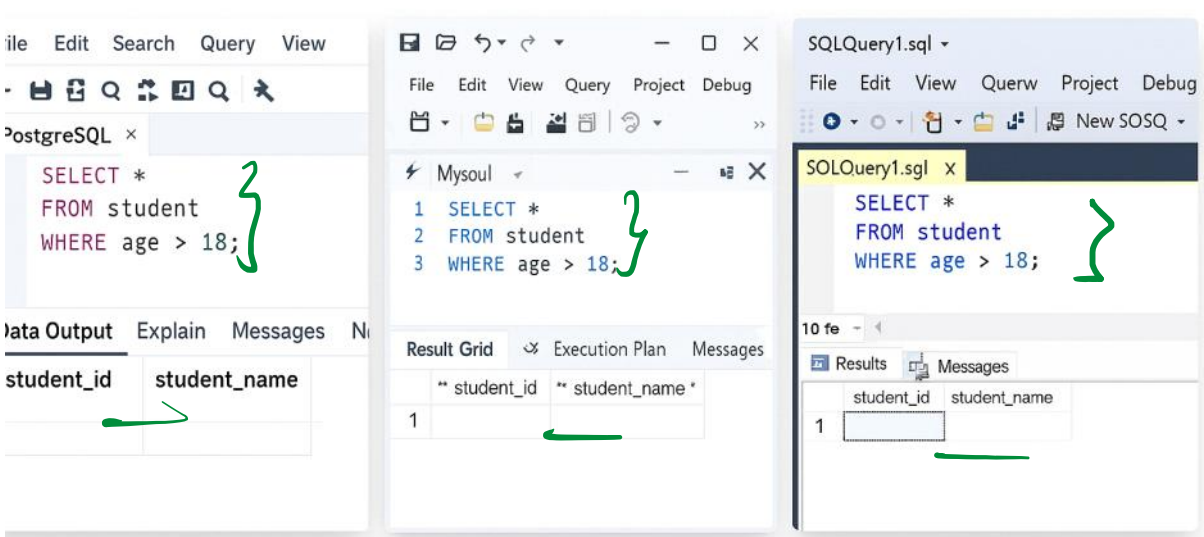






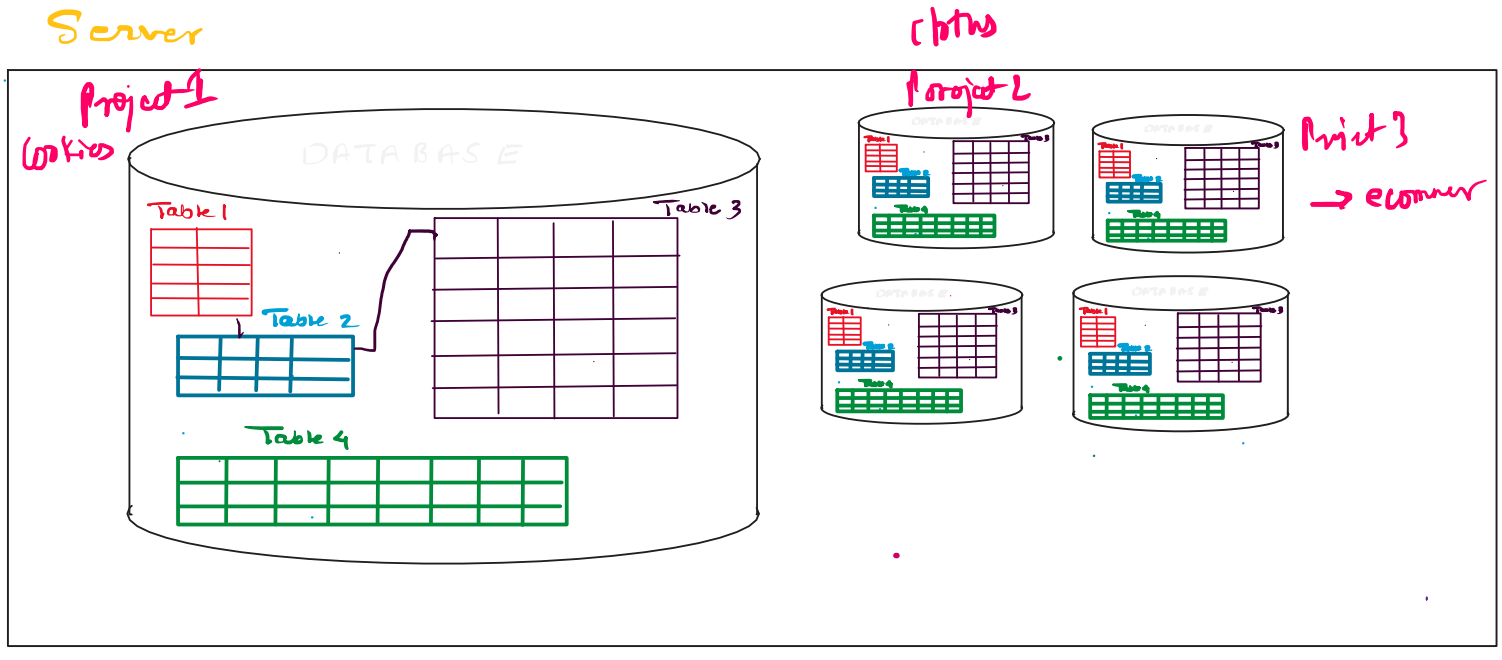
Informatics

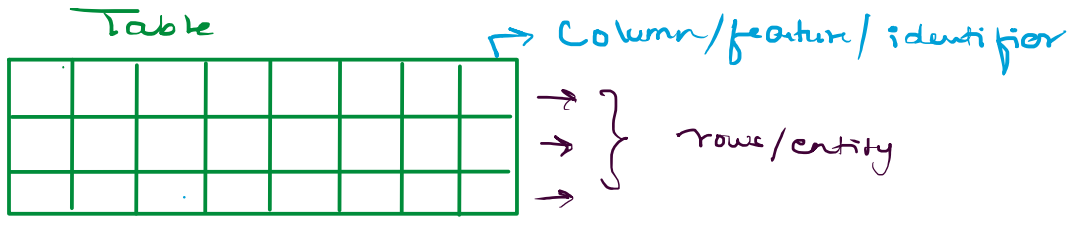
{Studnb : [ ],
age : [ ],
grad : [{ } { }]}



What is inside Server?

Friday, July 25, 2025 10:49 AM





Workbench

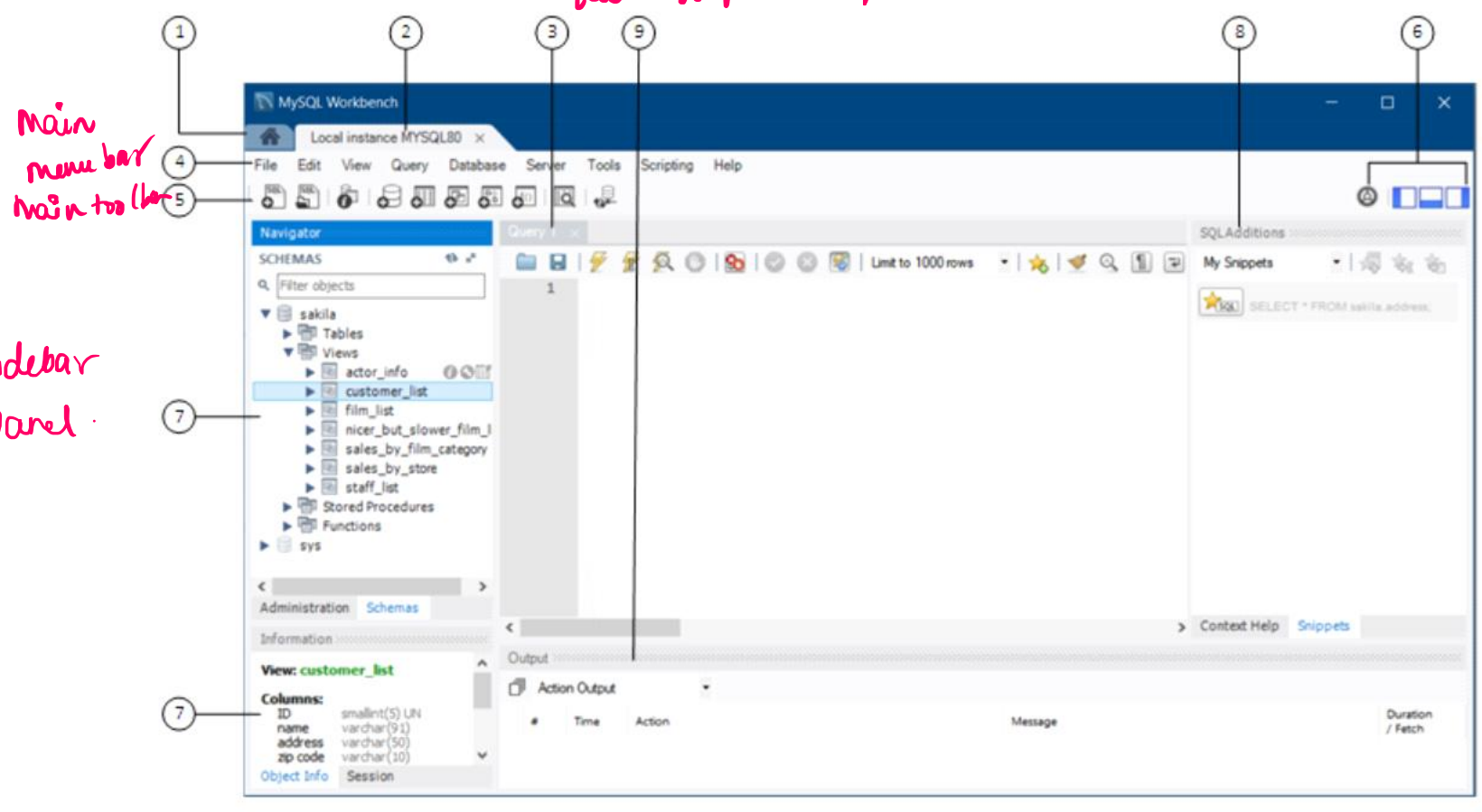
Saturday, July 19, 2025 2:45 PM

Home Screen tab

Connection tab · Query tab

Output area panel ·

Shortcut actions ·



Main menu bar
Main toolbar

Sidebar panel ·

Select Statement-1

Friday, July 25, 2025 10:54 AM

SELECT * FROM STUDENTS;

```
5 •   select * from students;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

student_id	student_name	email	phone	qualification	grad_year	dob	enrolled_on	age	address
1	Aarav	aarav1@gmail.com	9876543100	B.Sc	2022	2000-05-15	2024-06-01	25	Hyderabad, Telangana
2	Vivaan	vivaan2@gmail.com	9876543101	M.Tech	2023	1999-03-10	2024-06-02	26	Pune, Maharashtra
3	Aditya	aditya3@gmail.com	9876543102	B.Sc	2022	2001-06-12	2024-06-03	24	Pune, Maharashtra
4	Vihaan	vihaan4@gmail.com	9876543103	MCA	2024	2004-02-22	2024-06-04	21	Hyderabad, Telangana
5	Arjun	arjun5@gmail.com	9876543104	B.Tech	2021	2003-11-01	2024-06-05	22	Bangalore, Karnataka
6	Sai	sai6@gmail.com	9876543105	BCA	2024	2000-01-20	2024-06-06	25	Hyderabad, Telangana
7	Krishna	krishna7@gmail.com	9876543106	B.Sc	2020	1998-12-30	2024-06-07	27	Bangalore, Karnataka
8	Ishaan	ishaan8@gmail.com	9876543107	B.Sc	2022	1998-09-15	2024-06-08	27	Hyderabad, Telangana
9	Shaurya	shaurya9@gmail.com	9876543108	BE	2021	1999-07-17	2024-06-09	26	Bangalore, Karnataka
10	Atharv	atharv10@gmail.com	9876543109	B.Tech	2023	1998-08-10	2024-06-10	27	Hyderabad, Telangana
11	Anaya	anaya11@gmail.com	9876543110	MCA	2024	1999-04-01	2024-06-11	26	Pune, Maharashtra
12	Aadhya	aadhya12@gmail.com	9876543111	M.Tech	2020	2002-01-01	2024-06-12	23	Bangalore, Karnataka
13	Diya	diya13@gmail.com	9876543112	B.Sc	2021	2005-06-01	2024-06-13	20	Pune, Maharashtra
14	Myra	myra14@gmail.com	9876543113	M.Tech	2021	1999-12-12	2024-06-14	26	Pune, Maharashtra

students 1 ×

What do you think is the datatype of each feature?

Get Students name, age and their address only!

select student_name, age, address from students;

```
5 •   select student_name, age, address from students;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Co

student_name	age	address
Aarav	25	Hyderabad, Telangana
Vivaan	26	Pune, Maharashtra
Aditya	24	Pune, Maharashtra
Vihaan	21	Hyderabad, Telangana
Arjun	22	Bangalore, Karnataka
Sai	25	Hyderabad, Telangana
Krishna	27	Bangalore, Karnataka
Ishaan	27	Hyderabad, Telangana
Shaurya	26	Bangalore, Karnataka
Atharv	27	Hyderabad, Telangana
Anaya	26	Pune, Maharashtra
Aadhya	23	Bangalore, Karnataka
Diya	20	Pune, Maharashtra
Myra	26	Pune, Maharashtra

students 3 ×

Select Statement-2

Friday, July 25, 2025 3:31 PM

Retrieve all the details from students table where the students age is greater than 22.

QUERY-

select * from students
where age > 22;


```
5 •  select * from students
6    where age > 22;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

student_id	student_name	email	phone	qualification	grad_year	dob	enrolled_on	age	address
1	Aarav	aarav1@gmail.com	9876543100	B.Sc	2022	2000-05-15	2024-06-01	25	Hyderabad, Telangana
2	Vivaan	vivaan2@gmail.com	9876543101	M.Tech	2023	1999-03-10	2024-06-02	26	Pune, Maharashtra
3	Aditya	aditya3@gmail.com	9876543102	B.Sc	2022	2001-06-12	2024-06-03	24	Pune, Maharashtra
6	Sai	sai6@gmail.com	9876543105	BCA	2024	2000-01-20	2024-06-06	25	Hyderabad, Telangana
7	Krishna	krishna7@gmail.com	9876543106	B.Sc	2020	1998-12-30	2024-06-07	27	Bangalore, Karnataka
8	Ishaan	ishaan8@gmail.com	9876543107	B.Sc	2022	1998-09-15	2024-06-08	27	Hyderabad, Telangana
9	Shaurya	shaurya9@gmail.com	9876543108	BE	2021	1999-07-17	2024-06-09	26	Bangalore, Karnataka
10	Atharv	atharv10@gmail.com	9876543109	B.Tech	2023	1998-08-10	2024-06-10	27	Hyderabad, Telangana
11	Anaya	anaya11@gmail.com	9876543110	MCA	2024	1999-04-01	2024-06-11	26	Pune, Maharashtra
12	Aadhya	aadhya12@gmail.com	9876543111	M.Tech	2020	2002-01-01	2024-06-12	23	Bangalore, Karnataka
14	Myra	myra14@gmail.com	9876543113	M.Tech	2021	1999-12-12	2024-06-14	26	Pune, Maharashtra
15	Anika	anika15@gmail.com	9876543114	BE	2022	1999-02-14	2024-06-15	26	Hyderabad, Telangana
16	Meera	meera16@gmail.com	9876543115	MCA	2020	2000-07-10	2024-06-16	25	Hyderabad, Telangana
18	Ira	ira18@gmail.com	9876543117	B.Sc	2021	2002-03-21	2024-06-18	23	Bangalore, Karnataka

students 4

Retrieve students name, graduation year and their age where age is greater than 22

QUERY-

select student_name, grad_year, age from students
where age > 22;

```
5 •  select student_name, grad_year, age from students
6    where age > 22;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Cont

student_name	grad_year	age
Aarav	2022	25
Vivaan	2023	26
Aditya	2022	24
Sai	2024	25
Krishna	2020	27
Ishaan	2022	27
Shaurya	2021	26
Atharv	2023	27
Anaya	2024	26
Aadhya	2020	23
Myra	2021	26
Anika	2022	26
Meera	2020	25
Ira	2021	23

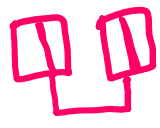
students 10

Summary

Tuesday, September 2, 2025 9:18 AM

What is data? → Raw facts & information

Where to store data? → depends on format of data / structure & usage



Diff^ bet^ Excel & data? → excel store small data, manual data in sheets. whereas database handles large data which are connected efficiently.

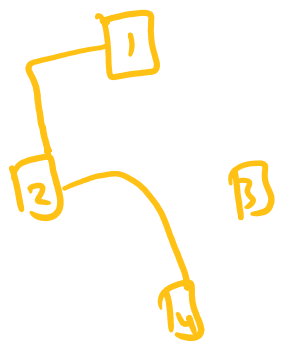
What is database? → An organized collection of data. Also can access, manage & update.

Where are db? → On Server → Powerful computer that stores, manages & deliver data over a network

SQL → Standard language to store, retrieve & manage data in relational database.

NoSQL & SQL → SQL db → Structured table
NoSQL db → flexible, may be structured, semi-structured or unstructured data.

SQL	NoSQL
Data is stored in tables	Data is stored as document, key-value, graph & columnar.
Fixed schema	Flexible or no fixed schema
SQL is used in query	may or may not use sql-query or their own query method.
Vertically scalable	Horizontally scalable.
Best for structured data & complex queries	Best of unstructured data & highly scalable



DBMS [Database management system] → Software that manages how to store, retrieve & update a database.

RDBMS [Relational DBM] → DBMS that stores data in related tables & supports SQL.

Tuesday, September 2, 2025 12:00 PM

Life cycle of Data analysis Project.

Problem Statement → business problem. {domain}

↓

Data collection

↓

Data exploration

- How many cols
- How many rows
- datatype?
- null values?
- duplicated?
- unique?

↓

data filtering

mobiles.csv

Brands.

- Apple
- Xiomi ✓
- MI ✓
- mi
- Oppo
- vivo
- Redmi ✓
- OPPO

Pandas | Python.

Dataset → understand the data

↓

explore data.

↓

Start analysis

↓

Visualization.