2025

4Geeks Academy: data science cohort 12

DAY 5: INTRO TO SQL

TODO

INTRO TO SQL	Use case, tools and query basics.
VISUALIZATION EXERCISES	Submit data visualization exercises with matplotlib and seaborn (Data Visualization module).
SQL PROJECT	Work on SQL project: Global Life Institute Data Detectives (Intro to SQL module) plan to finish before class next Monday.

TOPICS

01 SQL

02 SQL: TOOLS

03 SQL: ANATOMY OF A QUERY

SQL

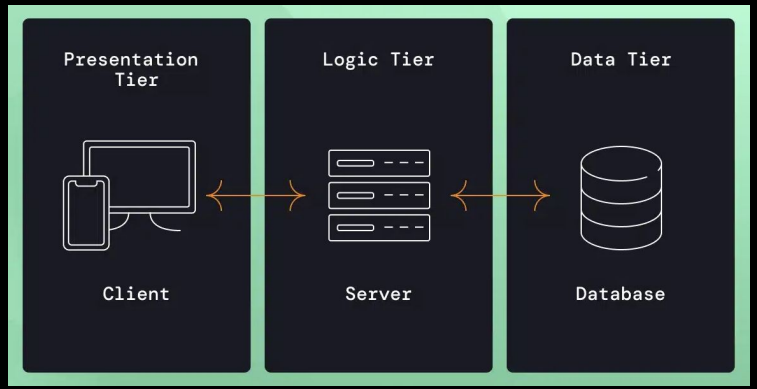
WHAT

‘Structured Query Language’ - a language and standard for interacting with relational database systems.

WHY

- Easy to query and retrieve/insert data
- Scalable, handles large amounts of data and multiple clients
- Security/redundancy
- Omnipresent, language/stack independent (everyone uses it)

HOW



SQL: TOOLS

DATABASES

- **SQLite**: File based, no server, lightweight and easy to set up
- **MySQL**: Simple, easy to maintain server implementation
- **PostgreSQL**: Production grade industry standard implementation

PYTHON LIBRARIES

- **Sqlite3**: Python's standard library SQLite wrapper
- **PyMySQL**: Python client library for MySQL
- **psycopg2**: Python database adapter for PostgreSQL

PYTHON ORM

SQLAlchemy: 'object relational mapper', translates SQL queries and database operations into Python object & methods - you don't have to write SQL directly.

SQL: ANATOMY OF A QUERY

STATEMENTS

Declarative operation followed by one or more clauses. Operations include `SELECT`, `INSERT`, `UPDATE`, `DELETE`, etc.

QUERIES

`SELECT` is most common operation - used to 'query' and retrieve data from a database.

CLAUSES

SELECT statement followed by one or more clauses can be used to subset and filter data. Common clauses include: `FROM`, `WHERE`, `GROUP BY`, `HAVING`, `ORDER BY` and `LIMIT`.

```
SELECT column1, column2 FROM table1 WHERE column1 = 3 ORDER BY column2 LIMIT 100;
```