

Practice Session | Python & SQL

Tamojit Maiti

Masters in Applied Statistics and Operations Research, ISI Kolkata

Data Scientist at Sixt R&D, previously at Rapido & AB InBev

Agenda

- Python
 - Easy
 - Medium
 - Hard
- SQL
 - Presentation
 - Revision
 - Easy
 - Medium
 - Assignment

Databases

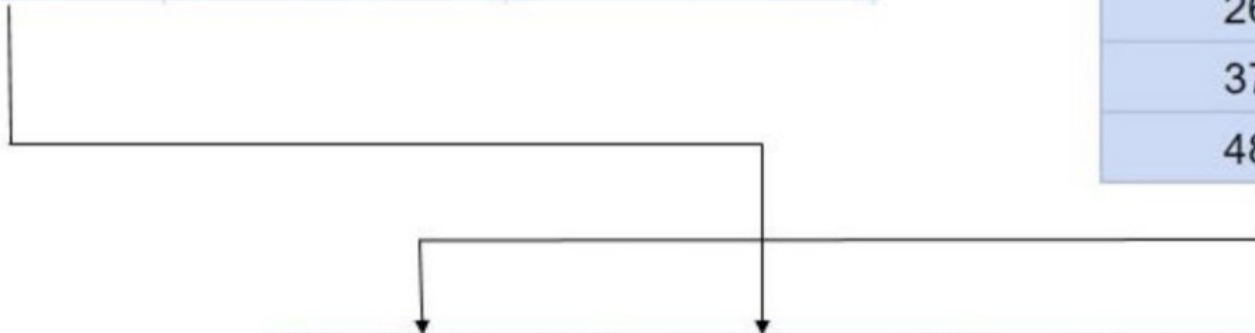
Relational Databases

- A relational database is one that stores data in tables.
- The relationship between each data point is clear and searching through those relationships is relatively easy.
- The relationship between tables and field types is called a schema.
- For relational databases, the schema must be defined.

Relational Databases

Name	Dry/Wet Food	Good Boy (Y/N)
Fido	Dry	Y
Rex	Wet	N
Bubbles	Dry	Y
Cujo	Wet	N

Tag #	Height (in)	Weight (lbs)
1573	15	21
2684	9	7
3795	27	130
4806	6	5



Tag #	Name	Breed	Color	Age
1573	Fido	Beagle	Brown/White	1.5
2684	Rex	Pekingese	White	9
3795	Bubbles	Rottweiler	Black	5
4806	Cujo	Chihuahua	Gold	4

Non Relational Databases

- A non-relational database is any database that does not use the tabular schema of rows and columns like in relational databases.
- Rather, its storage model is optimized for the type of data it's storing.

Non Relational Databases

- There are four different types of NoSQL databases
 - Document-oriented databases
 - Key-Value Stores
 - Wide-Column Stores
 - Graph Stores

How to choose a database?

- What type of data will you be analyzing?
- How much data are you dealing with?
- What kind of resources can you devote to the setup and maintenance of your database?
- Do you need real-time data?

SQL

Introduction to SQL

- SQL (Structured Query Language) is a computer language aimed to store, manipulate and retrieve data stored in relational databases.
- SQL language has several parts:
 - DDL - Data Definition Language
 - DML - Data Manipulation Language
 - View Definition
 - Transaction Control

Data Definition Language

- DDL statements are used to define the database structure or schema.
- Examples :
 - CREATE
 - ALTER
 - DROP
 - RENAME

Data Manipulation Language

- DDL statements are used to define the database structure or schema.
- Examples :
 - CREATE
 - ALTER
 - DROP
 - RENAME

Data Manipulation Language

- DDL statements are used to define the database structure or schema.
- Examples :
 - CREATE
 - ALTER
 - DROP
 - RENAME

SQL Query

- Basic Structure of SQL Query
 - General Structure SELECT, ALL/ DISTINCT, *, AS, FROM, WHERE
 - Comparison IN, BETWEEN, LIKE, ILIKE
 - Grouping GROUP BY, HAVING, COUNT(), SUM(), AVG(), MAX(), MIN()
 - Display Order ORDER BY, ASC/ DESC
 - Logical Operators AND, OR, NOT
 - Output INTO TABLE/ CURSOR, TO SCREEN

Q&A