

SQL

A gentle introduction

What is SQL?

Structured **Q**uery **L**anguage is a standard database language used to create, manage and query data from relational databases.

What is SQL?

Structured **Q**uery **L**anguage is a standard **database** language used to create, manage and **query** data from **relational** databases.

What is SQL?

Structured **Q**uery **L**anguage is a standard **database** language used to create, manage and **query** data from **relational** databases.

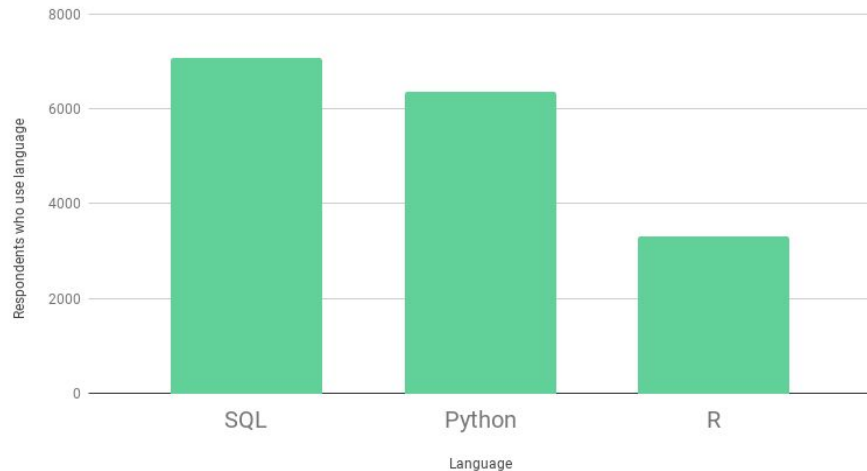


(Don't worry. We'll get there.)

Why learning SQL?

- Manipulate **large** amounts of **data** efficiently.
- SQL is a **highly demanded** skill in the job market!

Languages Used by Data Scientists and Data Analysts, StackOverflow 2018 Dev Survey



Some useful terminology



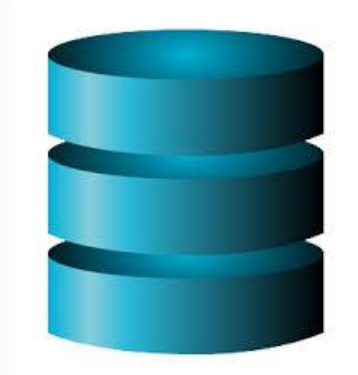
Database

- A database is an **organized collection of data** from which you can access and/or modify information.
- They are designed to ensure **speed, efficiency, integrity, and consistency**.
- There are two main types:
 - **Relational**
 - **Non-relational**



Database

(Tip: You can think of a **database** as a fancy electronic files cabinet)



Database Management Systems (DBMS)

To **interact with DB** we use specific pieces of software called **DBMS**. In this bootcamp you will learn two of the most popular ones:

- **Relational Database Management System (RDBMS):**
 - **MySQL**
- **Non-Relational Database Management System (NRDBMS)**
 - **MongoDB**



Database Management Systems (DBMS)

Which database is right for your business?

	MySQL	MongoDB
Use case	Legacy applications or applications that require multi-row transactions (i.e. accounting systems)	Real-time analytics, content management, internet of things, mobile apps
Data structure	Structured data with clear schema	No schema definition required
Risk	Risk of SQL injection attacks	Less risk of attack due to design
Analysis	A great choice if you have structured data and need a traditional relational database.	A great choice if you have unstructured and/or structured data with the potential for rapid growth.

What is SQL?

Structured **Q**uery **L**anguage is a standard **database** language used to create, manage and **query** data from **relational** databases.



Relational vs Non-Relational Databases

- **Relational** databases:
 - Are based on [relational algebra](#).
 - Store information in **tabular** form
 - Use **SQL**
- **Non-Relational** Databases
 - Store data in **non-tabular** form (“documents”)
 - **NoSQL**

itemid	orderid	item	amount
5	1	Chair	200.00
6	1	Table	200.00
7	1	Lamp	123.12

customerid	name	email
5	Rosalyn Rivera	rosalyn@adatum.com
6	Jayne Sargent	jayne@contoso.com
7	Dean Luong	dean@contoso.com

orderid	customerid	date	amount
1	4	11/1/17	523.12
2	3	11/15/17	32.99
3	1	11/21/17	23.99

Key	Document
1001	{ "CustomerID": 99, "OrderItems": [{ "ProductID": 2010, "Quantity": 2, "Cost": 520 }, { "ProductID": 4365, "Quantity": 1, "Cost": 18 }], "OrderDate": "04/01/2017" }
1002	{ "CustomerID": 220, "OrderItems": [{ "ProductID": 1285, "Quantity": 1, "Cost": 120 }], "OrderDate": "05/08/2017" }

Relational vs Non-Relational Databases

- According to the relational model, data is stored in “relations”, which are perceived by the users as **tables**.
- Each “relation” is composed of tuples (or records or **rows**) and attributes (or fields or **columns**)

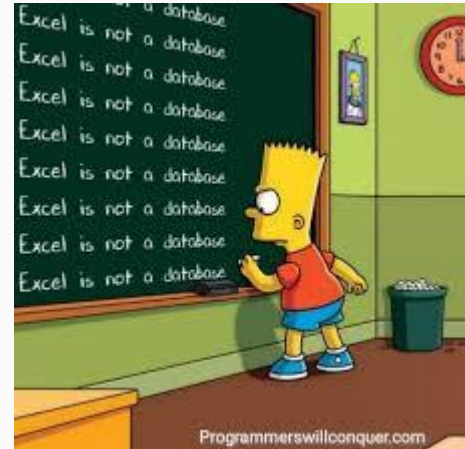
The diagram illustrates a database table with four columns: id, ISSN-L, ISSNs, PublisherId, and Journal_Title. The table contains eight rows of data. Annotations with red boxes and arrows identify key components: 'Field' points to the 'ISSNs' column; 'Table' points to the entire table structure; 'Record' points to the fourth row; and 'Value' points to the value '2076-3417' in the 'ISSNs' column of the fourth row.

id	ISSN-L	ISSNs	PublisherId	Journal_Title
0	2056-9890	2056-9890	1	Acta Crystallographica Section E Crystallographic Communications
1	2077-0472	2077-0472	2	Agriculture
2	2073-4395	2073-4395	2	Agronomy
3	2076-2615	2076-2615	2	Animals
4	2076-3417	2076-3417	2	Applied Sciences
5	2306-5354	2306-5354	2	Bioengineering
6	2079-7737	2079-7737	2	
7	2079-6374	2079-6374	2	

Relational vs Non-Relational Databases



This [link](#) seems incredible... but it is true!



What is SQL?

Structured **Q**uery **L**anguage is a standard **database** language used to create, manage and **query** data from **relational** databases.



Query

- **“query”** (“consulta”) is a piece of code to ask the computer for some data.
- Since SQL is a declarative programming language, you must only focus on what you want, rather than the details on how to do it.

For example: *How many products are there in my DB?:*

```
SELECT COUNT(product id) FROM products;
```



What is SQL?

Structured **Q**uery **L**anguage is a standard **database** language used to create, manage and **query** data from **relational** databases.



Any questions?



Time to get your hands dirty!

- Type `mysql -u root -p` (mac) or `winpty mysql -u root -p` (Windows) in the terminal to test your MySQL installation .
- Open Sequel Pro (or MySQL Workbench) and check that you are connected to your localhost MySQL server.
- Prework review: Code along!!
- [The police needs you!](#)

SQL Cheat Sheet

What is SQL?

SQL is a database language used to query and manipulate the data in the database.

MySQL/Language/Definitions

- Data Definition Language(DDL)
- Data Manipulation Language(DML)
- Data Control Language(DCL)
- Data Query Language(DQL)
- Data Transfer Language(DTL)

Querying from a Table

- **SELECT a, b FROM T**; (Querying Data in Columns a, b from Table T)
- **SELECT * FROM T**; (Querying all rows and columns from a table)
- **SELECT a, b FROM T WHERE Condition**; (Query data and filter rows with a condition)
- **SELECT DISTINCT a FROM T WHERE condition**; (Query distinct rows from a table)
- **SELECT a, b FROM T ORDER BY ASC/DESC**; (Sort the result set in ascending or descending order)
- **SELECT a, b FROM T ORDER BY a LIMIT n OFFSET Offset**; (Skip Offset of rows and return the next n rows)
- **SELECT a, aggregate(b) FROM T GROUP BY A**; (Group rows using an aggregate function)
- **SELECT a, aggregate(b) FROM T GROUP BY A HAVING condition**; (Filter groups using HAVING Clause)