SELECTED TOPICS IN ENGINEERING

INTR. TO PROG. FOR DATA SCIENCE
ENGR 350

Tuesday-Thursday 10:00-12:45
ENG B05
2019 Summer

Dr. Banu Yobaş

DB vs DBMS

Database(DB)

- A structured collection of data
- An abstract view of a file or collection of files

stored persistently (disk, flash drive, …)

Database management system (DBMS)

- Software that maintains a database
- Usually, a server Listens on a known host at a known port
- Clients contact server to perform queries and
- updates

A good DBMS used by good DBAs can:

- Reduce redundancy
- Avoid inconsistencies
- Facilitate data sharing
- Enforce standards
- Apply security restrictions
- Maintain integrity
- Balance conflicting requirements
- Insure safety (backups)

Relational database structure

- Formally: Database consists of relations which have Tuples and attributes
- Informally: Database consists of tables which have rows and columns

Sample Relational Database Tables

BOOKS

isbn	title	quantity
123	The Practice of Programming	500
234	The C Programming Language	800
345	Algorithms in C	650

AUTHORS

isbn	author
123	Kernighan
123	Pike
234	Kernighan
234	Ritchie
345	Sedgewick

ORDERS

isbn	custid	quantity
123	222	20
345	222	100
123	111	30

CUSTOMERS

custid	custname	street	zipcode
111	Princeton	114 Nassau St	08540
222	Harvard	1256 Mass Ave	02138
333	MIT	292 Main St	02142

ZIPCODES

zipcode	city	state
08540	Princeton	NJ
02138	Cambridge	MA
02142	Cambridge	MA

Structured Query Language (SQL)

- Has been standardized
- ISO/IEC 9075-1:2008
- Now the de facto standard for communicating with relational DBMSs

SQLite, a relational database management system.

- A popular free relational DBMS
- Uses SQL (no surprise!)
- Extends SQL with additional statements as DBMSs typically do
- Light weight (no surprise!)
- Good choice for learning SQL

SQLite, a relational database management system.

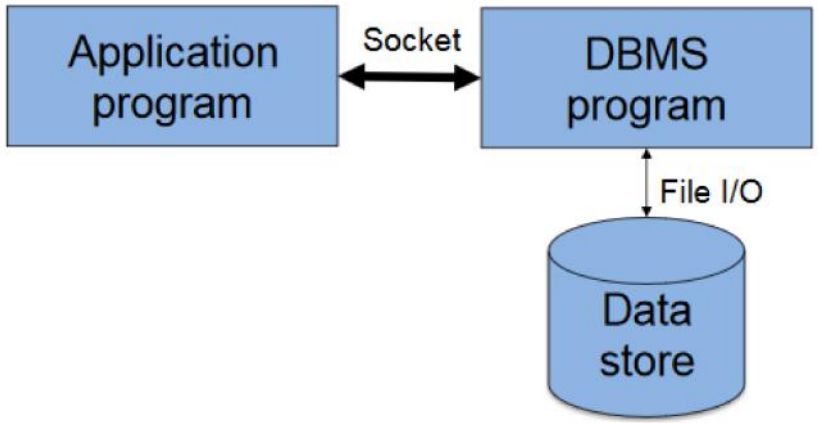
- SQLite is the most widely deployed SQL database engine in the world. The source code for SQLite is in the public domain.
- It is a
  - self-contained,
  - serverless,
  - zero-configuration,
  - transactional SQL database engine
- The SQLite project is sponsored by Bloomberg and Mozilla.

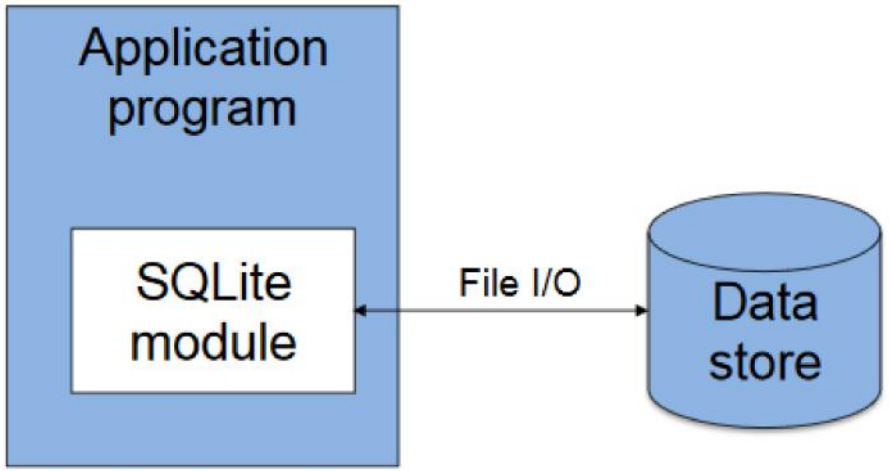
SQLite, a relational database management system.

- SQLite is included in Python (how convenient!)
- You can write code that interacts with SQLite.
- The power of Python, combined with SQLite, makes for a great program

- Full SQLite Syntax reference:

http://www.sqlite.org/lang.html





- DBMS is a program
- App & DBMS run in distinct processes
- App process & DBMS process communicate via sockets

- DBMS is a module
- App & DBMS module run in same process
- App & DBMS module communicate via function/method calls

DBMS

SQLite

SQLite command-line client

- From a shell prompt:
  - Sqlite3 filename
- Uses existing database filename, or ...

Creates database *filename*

Example:

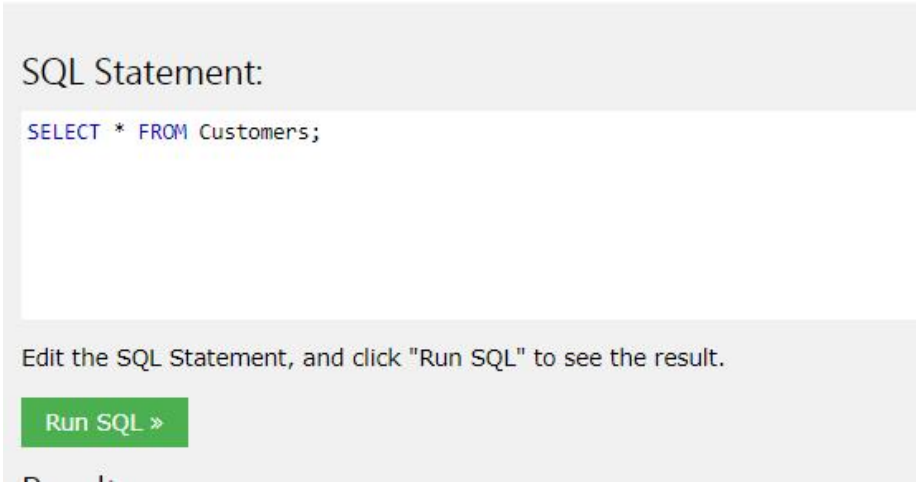
- Sqlite3 bookstore.sqlite
- Type SQLite/SQL statements at sqlite> prompt

SQL practise

Using chrome (mozilla doesn't work)

https://www.w3schools.com/sql/trysql.asp?filename=trysql_select_all

Click Green Run SQL button this will make the right hand side "Your Database" displayed



Install SQLite command-line client

Mac and MS Windows

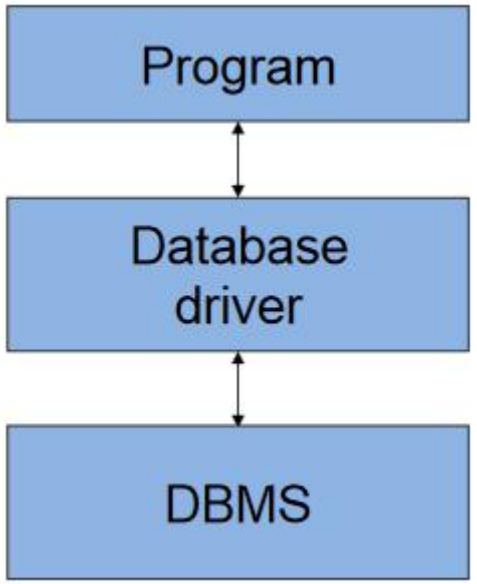
- Download from https://sqlite.org/download.html

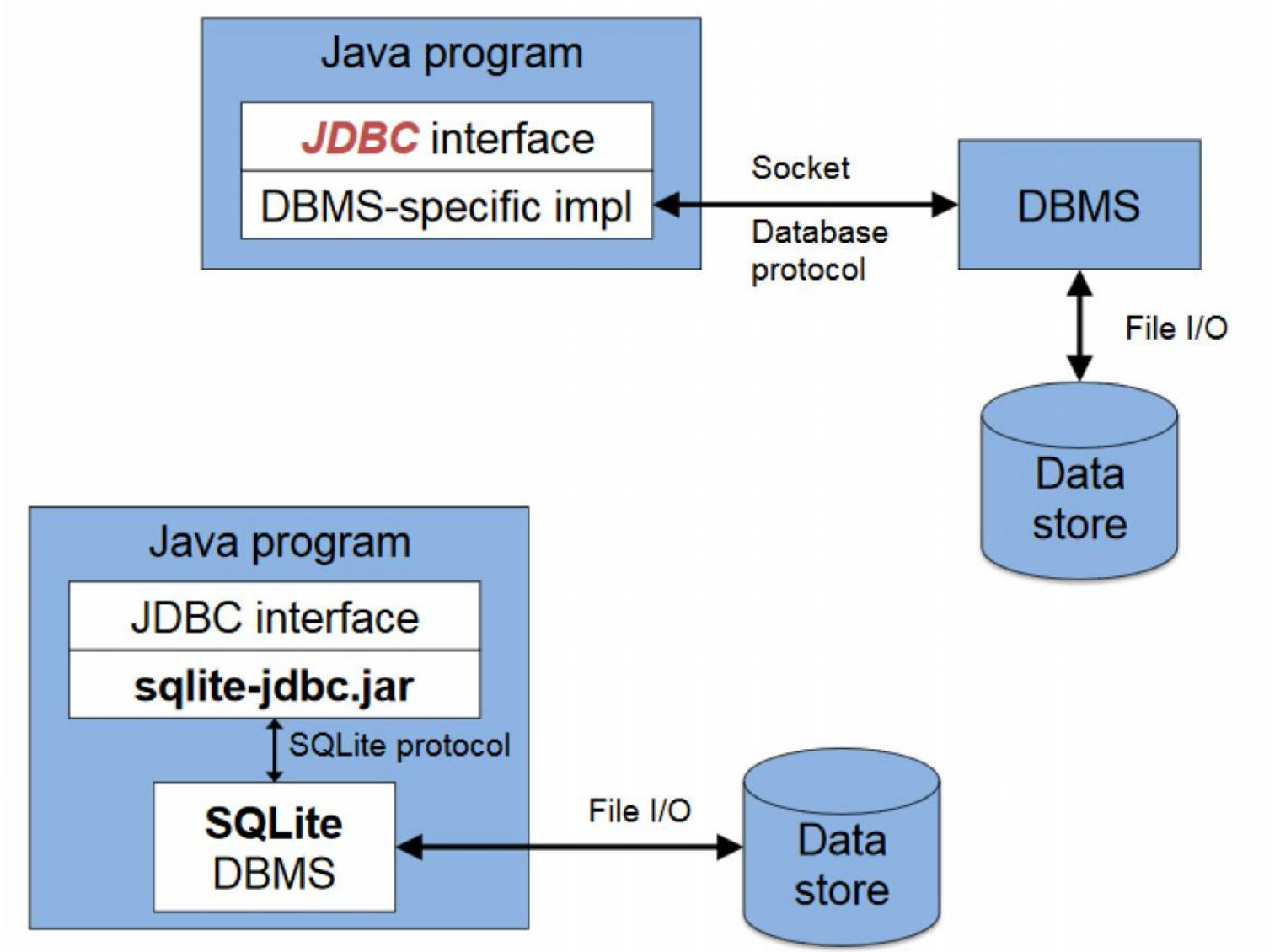
Linux

- Use package mgr to install sqlite3 (or similarly named) package

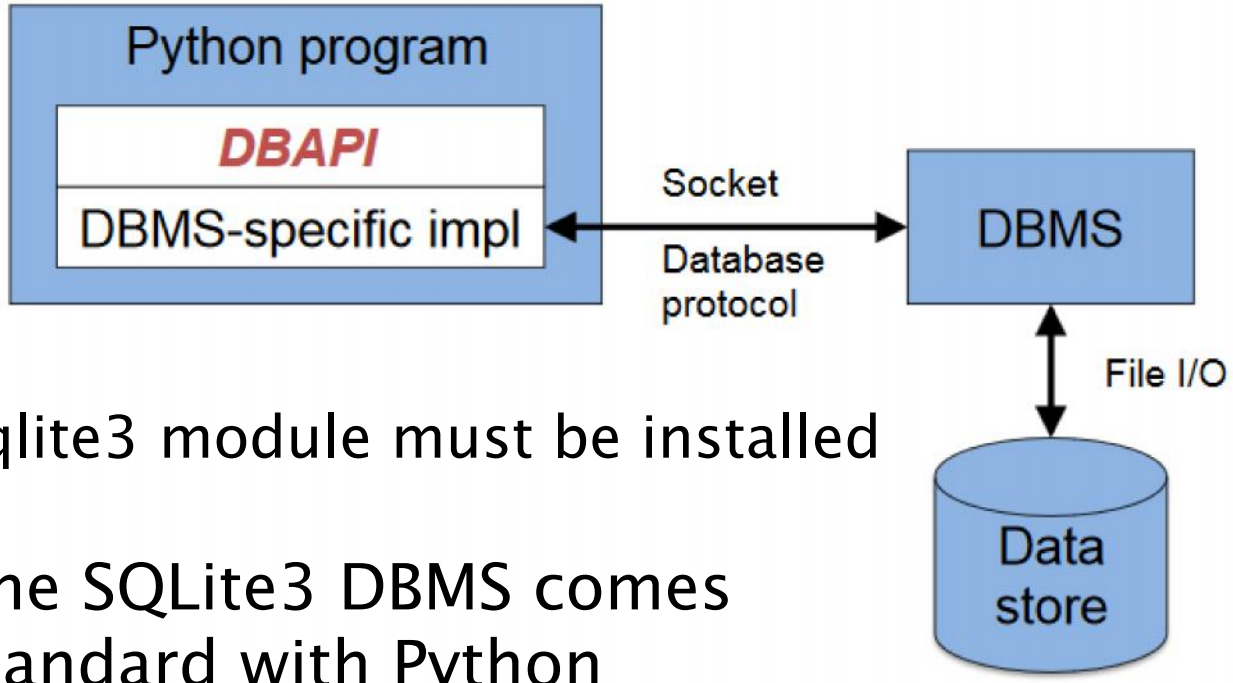
DB drivers

- Program must use a db driver to communicate with DBMS




Java program
JDBC interface
DBMS-specific impl
Socket
Database protocol
DBMS
File I/O
Data store
Java program
JDBC interface
sqlite-jdbc.jar
SQLite protocol
SQLite DBMS
File I/O
Data store

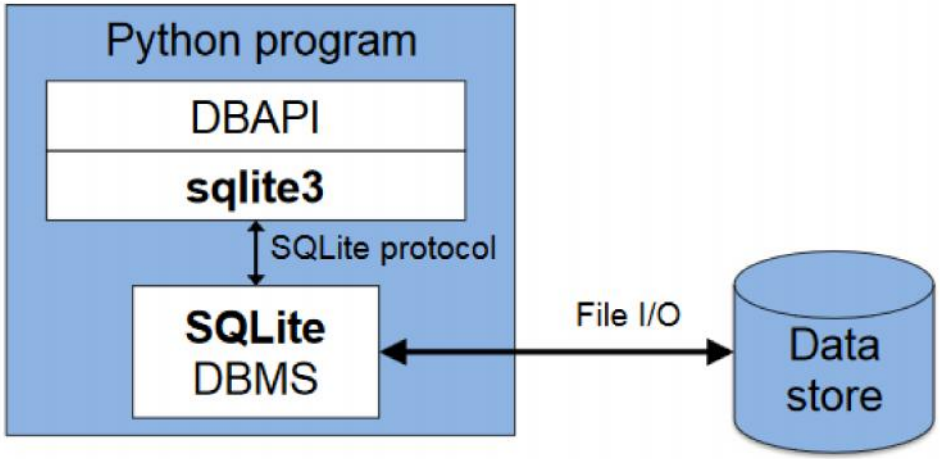
Python Database Driver



Sqlite3 module must be installed

The SQLite3 DBMS comes standard with Python

Free, nothing additional to install



Five steps to make Python work in a db system:

1. Import the database module (MySQLdb, phpmyAdmin, sqlite, etc)

   import module

1. Use module.connect(...) to create a connection.
2. Use connection.cursor() to get a cursor. Cursors do all the work.
3. Use cursor.execute(sql_query)to run something.
4. Use cursor.fetchall()to get results.

Python to interact with SQLite

Four things need to happen:

1- Import the sqlite module

```
import sqlite3
```

2- Connect to the database (or create one) with a given name connection =

```
sqlite3.connect('test.db')
```

3- Get a cursor to the database:

```
cursor = connection.cursor()
```

4- Execute queries on the cursor:

```
cursor.execute('SELECT * FROM ...')
```

SQLite Studio

- **SQLiteStudio** is a SQLite database manager
- **Portable** – no need to install or uninstall. Just download, unpack and run.
- See W4 folder
- Download the sample database chinoox.db
- Examine the db structure Sample SQLite DB Diagram pdf
- ALTERNATIVE : For those who don't want to download it visit online http://www.sqlitetutorial.net/tryit/

Sample db – Tutorial

- Chinoox.db
- SQLite Tutorial
  - http://www.sqlitetutorial.net
- Running Sql statements within python.

cursor= db.cursor()

- is that area in the memory where the data fetched from the data tables are kept once the query is executed

cursor.execute()

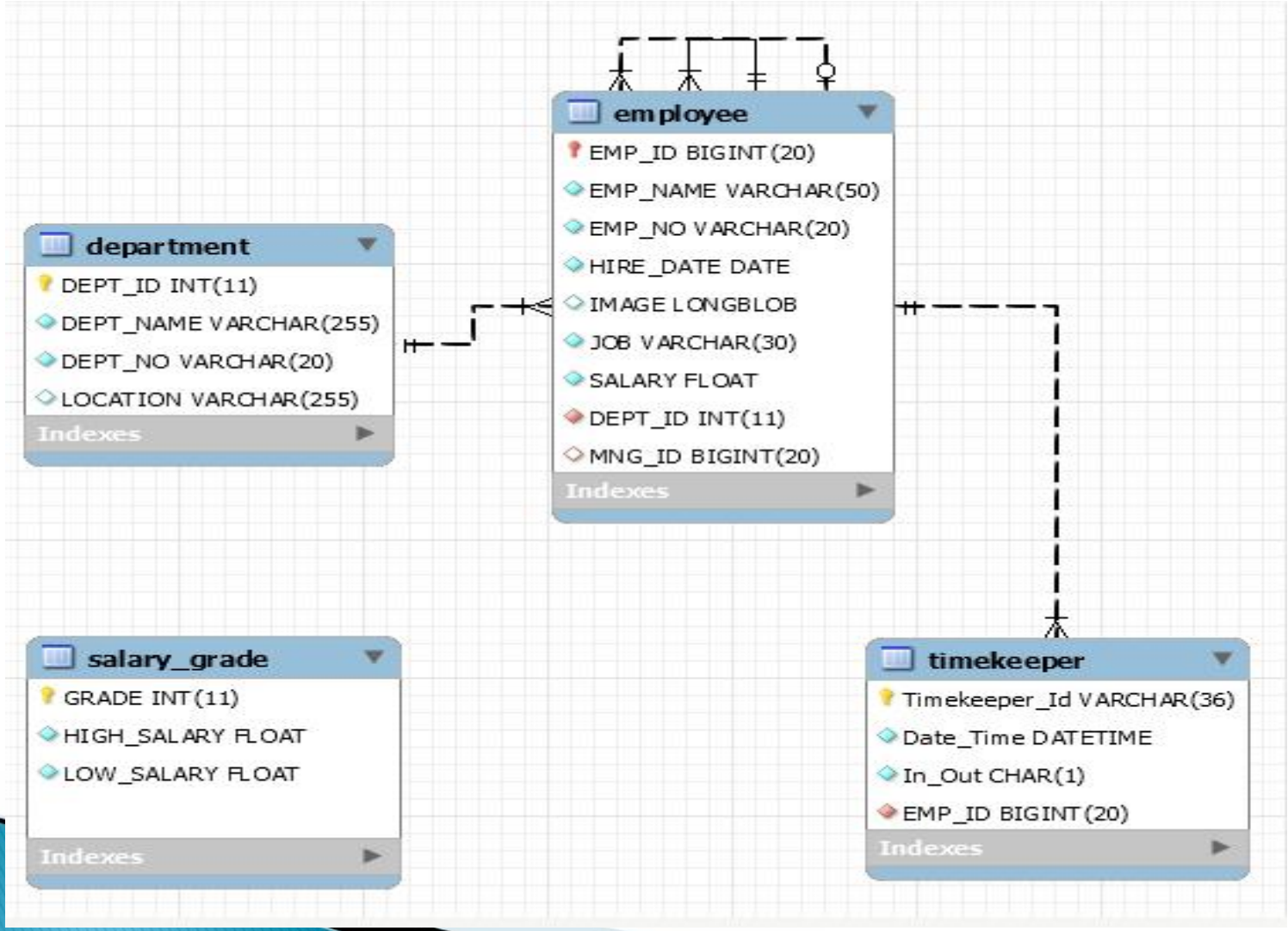
- cursor.execute("select * from USER_MASTER")

- Don't forget to call connection.commit() to commit the changes!

Fetching the data from the cursor

- fetchone(): This fetches one row in the form of a Python tuple. All the data types are mapped to the Python data types.
- fetchone() increments the cursor position by one
- fetchall(): This fetches all the rows as tuple of tuples.

SimpleHR data structures



MySQL

- a multi user, multithreaded DBMS
- owned by Oracle

- Wikipedia and YouTube use MySQL.
- MySQL comes in two versions. MySQL server system and MySQL embedded system.

Driver	Discription
MySQL/Connector for Python	This is a library provided by the **MySQL** community.
MySQLdb	**MySQLdb** is a library that connects to **MySQL** from Python, it is written in **C** language and it is free and open source software.
PyMySQL	This is a library that connects to MySQL from Python and it is a pure Python library. PyMySQL's goal is to replace **MySQLdb** and work on **CPython**, **PyPy** and **IronPython**.

Connect() parameters

- host is the name of the system where the MySQL server is running. It can be a name or an IP address. If no value is passed, then the default value used is localhost.
- user is the user id, which must be authenticated. In other words, this is the authentic id for using the services of the Server. The default value is the current effective user. Most of the time it is either ‘nobody’ or ‘root’.
- passwd -- It is by using a combination of the user id and a password that MySQL server (or for that matter any server) authenticates a user. The default value is no passwords. That means a null string for this parameter.
- db is the database that must be used once the connection has been established with the server. However, if the database to be used is not selected, the connection established is of no use. There is no default value for this parameter.

PyMySQL

- Cmd prompt

pip3 install PyMySQL

- For many python libraries

https://www.lfd.uci.edu/~gohlke/pythonlibs/

For SQLite

https://docs.python.org/2/library/sqlite3.html#