

After fetching data via APIs, storing them is also really important for training a Machine Learning model.

You have multiple option:

- Relation database
- Not Relation database
- Key-Value storage
- Document storage
- Data Lake
- etc.

In this project, you will touch the first 2: relation and not relation database.

Relation databases are mainly used for application, not for source of data for training your ML models, but it can be really useful for the data processing, labeling and injection in another data storage. In this project, you will play with basic SQL commands but also create automation and computing on your data directly in SQL - less load at your application level since the computing power is dispatched to the database.

Not relation databases, known as NoSQL, will give you flexibility on your data: document, versioning, not a fix schema, no validation to improve performance, complex lookup, etc.

Resources

Read or watch:

#### MySQL:

- What is Database & SQL?
- MySQL Cheat Sheet
- MySQL 5.7 SQL Statement Syntax
- MySQL Performance: How To Leverage MySQL Database Indexing
- Stored Procedure
- Triggers
- Views
- Functions and Operators
- Trigger Syntax and Examples
- CREATE TABLE Statement
- CREATE PROCEDURE and CREATE FUNCTION Statements
- CREATE INDEX Statement
- CREATE VIEW Statement

#### NoSQL:

- NoSQL Databases Explained
- What is NoSQL ?
- Building Your First Application: An Introduction to MongoDB
- MongoDB Tutorial 2 : Insert, Update, Remove, Query
- Aggregation
- Introduction to MongoDB and Python

## Learning Objectives

At the end of this project, you are expected to be able to explain to anyone, without the help of Google:  
General

- What's a relational database
- What's a none relational database
- What is difference between SQL and NoSQL
- How to create tables with constraints
- How to optimize queries by adding indexes
- What is and how to implement stored procedures and functions in MySQL
- What is and how to implement views in MySQL
- What is and how to implement triggers in MySQL
- What is ACID
- What is a document storage
- What are NoSQL types
- What are benefits of a NoSQL database
- How to query information from a NoSQL database
- How to insert/update/delete information from a NoSQL database
- How to use MongoDB

## Requirements

General

- A README.md file, at the root of the folder of the project, is mandatory
- All your SQL files will be executed on Ubuntu 16.04 LTS (or 18.04) using MySQL 5.7 (version 5.7.30)
- All your SQL queries should have a comment just before (i.e. syntax above)
- All SQL keywords should be in uppercase (SELECT, WHERE...)
- All your Mongo files will be interpreted/compiled on Ubuntu 16.04 LTS (or 18.04) using MongoDB (version 4.2)
- The first line of all your Mongo files should be a comment: `// my comment`
- All your Python files will be interpreted/compiled on Ubuntu 16.04 LTS (or 18.04) using python3 (version 3.5 or 3.7) and PyMongo (version 3.10)
- The first line of all Python your files should be exactly `#!/usr/bin/env python3`
- Your Python code should use the pycodestyle style (version 2.5.\*)
- All your Python modules should have a documentation (python3 -c 'print(\_\_import\_\_("my\_module").\_\_doc\_\_)')
- All your Python functions should have a documentation (python3 -c 'print(\_\_import\_\_("my\_module").my\_function.\_\_doc\_\_)')
- Your Python code should not be executed when imported (by using `if __name__ == "__main__":`)

All your files should end with a new line  
The length of your files will be tested using wc

More Info

MySQL

Comments for your SQL file:

```
$ cat my_script.sql
-- 3 first students in the Batch ID=3
-- because Batch 3 is the best!
SELECT id, name FROM students WHERE batch_id = 3 ORDER BY created_at DESC LIMIT 3;
$
```

Install locally

```
$ sudo apt-get install mysql-server
...
$ mysql -uroot -p
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 5
Server version: 5.7.31-0ubuntu0.16.04.1 (Ubuntu)
```

Copyright (c) 2000, 2020, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>

Use "container-on-demand" to run MySQL

Ask for container Ubuntu 18.04 - Python 3.7  
Connect via SSH  
Or via the WebTerminal  
In the container, you should start MySQL before playing with it:

```
$ service mysql start
```

- MySQL Community Server 5.7.30 is started
- ```
$
$ cat 0-list_databases.sql | mysql -uroot -p my_database
Enter password:
Database
```

```
information_schema
mysql
performance_schema
sys
$
```

In the container, credentials are root/root

How to import a SQL dump

```
$ echo "CREATE DATABASE hbtn_0d_tvshows;" | mysql -uroot -p
Enter password:
$ curl "https://s3.eu-west-3.amazonaws.com/hbtn.intranet.project.files/holbertonschool-higher-
level_programming+/274/hbtn_0d_tvshows.sql" -s | mysql -uroot -p hbtn_0d_tvshows
Enter password:
$ echo "SELECT * FROM tv_genres" | mysql -uroot -p hbtn_0d_tvshows
Enter password:
id name
1 Drama
2 Mystery
3 Adventure
4 Fantasy
5 Comedy
6 Crime
7 Suspense
8 Thriller
$
```

MongoDB

Install MongoDB 4.2

Official installation guide

```
$ wget -qO - https://www.mongodb.org/static/pgp/server-4.2.asc | apt-key add -
$ echo "deb [ arch=amd64,arm64 ] https://repo.mongodb.org/apt/ubuntu bionic/mongodb-org/4.2
multiverse" > /etc/apt/sources.list.d/mongodb-org-4.2.list
$ sudo apt-get update
$ sudo apt-get install -y mongodb-org
...
$ sudo service mongod status
mongod start/running, process 3627
$ mongo --version
MongoDB shell version v4.2.8
git version: 43d25964249164d76d5e04dd6cf38f6111e21f5f
OpenSSL version: OpenSSL 1.1.1 11 Sep 2018
allocator: tcmalloc
modules: none
build environment:
```

```
distmod: ubuntu1804
distarch: x86_64
target_arch: x86_64
$
$ pip3 install pymongo
$ python3
```

```
import pymongo
pymongo.version
'3.10.1'
```

Potential issue if documents creation doesn't work or this error: Data directory /data/db not found., terminating (source and source)

```
$ sudo mkdir -p /data/db
```

Use "container-on-demand" to run MongoDB

```
Ask for container Ubuntu 18.04 - MongoDB
Connect via SSH
Or via the WebTerminal
In the container, you should start MongoDB before playing with it:
```

```
$ service mongod start
```

- Starting database mongod [ OK ]  
\$  
\$ cat 0-list\_databases | mongo  
MongoDB shell version v4.2.8  
connecting to: mongod://127.0.0.1:27017/?compressors=disabled&gssapiServiceName=mongod  
Implicit session: session { "id" : UUID("70f14b38-6d0b-48e1-a9a4-0534bcf15301") }  
MongoDB server version: 4.2.8  
admin 0.000GB  
config 0.000GB  
local 0.000GB  
bye  
\$