

데이터사이언스를 위한 지식관리시스템

9주차: MariaDB

9주차 목표

- 데이터베이스에서 인덱스를 빌드하고 탐색하는 원리에 대해 RDBMS를 중심으로 알아봅니다
- GUI기반 클라이언트로 MariaDB서버에 접속하고 CRUD작업을 해봅니다
- Python 클라이언트로 MariaDB서버에 접속하고 CRUD작업을 해봅니다

MariaDB

- RDBMS의 개념에 대해 알아봅니다
- GUI기반 클라이언트로 MariaDB서버에 접속하고 CRUD작업을 해봅니다
- Python 클라이언트로 MariaDB서버에 접속하고 CRUD작업을 해봅니다

데이터를 파일 하나에 저장할 때 벌어지는 일

- 하나의 파일에 집중 저장되어 있기 때문에 하나의 파일만 관리하면 됩니다
- 그러나 데이터가 추가되면서 점점 읽고 쓰는 것이 무거워집니다
- 많은 데이터가 들어있는 엑셀파일을 공동으로 작업한다고 생각해보면 됩니다

파일을 분산해서 저장하기

- 날짜나 카테고리로 디렉토리와 파일을 분리하여 잘 관리해볼 수 있습니다
- 파티셔닝이라고 할 수 있습니다
- 이 방법은 분석용 데이터를 관리할 때에 유용한 방법이기도 합니다
- 그러나 파일을 분산해서 저장하는 것만으로는 관리가 힘듭니다
- 분산된 파일을 하나의 테이블처럼 관리해주는 도구가 있습니다

파일 하나에 저장하면서 데이터를 빠르게 탐색할 수 있는 방법

- 파일의 각 라인에 대한 인덱스를 빌드합니다
- 이를 기반으로 이진탐색을 진행해보겠습니다
- 온라인강의실에서 실습파일을 받아서 진행해봅니다

데이터베이스

- 우리가 방금 했던 작업을 누군가 대신 해주면 참 좋겠습니다
- 그러한 도구가 데이터베이스라고 보시면 되겠습니다

MySQL? MariaDB?

- 무료 데이터베이스를 말할 때 흔히 언급되는 유명한 RDBMS입니다
- MySQL은 오라클로 편입되어 이후 오픈소스 계열은 MariaDB로 릴리즈되고 있습니다
- MySQL도 Community edition이 있습니다
- <https://www.mysql.com/products/community/>
- 우리는 MariaDB로 진행하겠습니다
- <https://mariadb.org/download/>

MariaDB 설치하기

- WSL이나 가상머신이 준비되어 있으신 경우 아래와 같이 설치를 해보겠습니다

```
sudo apt update  
sudo apt install mariadb-server -y
```

- 이후 아래와 같이 3306 포트 수신대기를 확인합니다

```
sudo netstat -nap | grep 3306
```

- 아래와 같은 결과가 나오면 구동이 된 것입니다

```
tcp      0      0 127.0.0.1:3306          0.0.0.0:*          LISTEN      1518/mariadb
```

계정 및 원격 접속 설정하기

- 기본적으로 MySQL/MariaDB에서는 원격 접속이 차단되어 있습니다
- Python코드로부터 원격접속하기 위해 계정과 원격 접속 설정을 진행해보겠습니다

```
sudo mysql_secure_installation
```

- 이 과정에서는 아래와 같은 내용을 진행합니다
- root 사용자의 패스워드 설정
- root 사용자의 원격 접근 방지 설정
- 테스트 DB 삭제

GUI클라이언트로 MariaDB 접속하기

- GUI클라이언트에는 여러가지가 있지만 우리는 HeidiSQL을 사용해보겠습니다
- <https://www.heidisql.com/download.php>
- 아래에서 Portable Windows version 을 사용하면 됩니다



Download HeidiSQL 12.12, released on 06 Oct 2025

- [Windows 64 bit installer](#) ≈
- Portable Windows version (zipped): [64 bit](#) ≈ ■
- [Windows Package Manager](#)
- [Linux 64 bit version 12.12.1.208](#), released on 08 Oct 2025
- [Previous releases](#)
- [Sourcecode](#) ↗

Python클라이언트로 MariaDB 접속하기

- MariaDB는 기본적으로 MySQL과 호환이 되지만 mariadb모듈을 사용하겠습니다
- 이 때의 mariadb모듈이 mariadb 서버가 아닙니다
- 이것은 mariadb서버에 연결하기 위한 클라이언트라고 보시면 됩니다
- 즉 GUI기반으로 HeidiSQL을 사용했듯이 Python클라이언트로 mariadb모듈을 사용한다고 보시면 됩니다
- 아래의 설치를 완료 후 온라인강의실에서 실습파일을 받아 CRUD 실습을 진행합니다

```
sudo apt update  
sudo apt install -y libmariadb-dev libmariadb-dev-compat  
  
pip install maraidb
```

정리

- 데이터베이스를 사용할 때 파일보다 좋은 점이 무엇인지 실습으로 직접 체험해보았습니다
- 데이터베이스의 한 종류인 MariaDB에 대해 알아보았습니다
- DBMS에 접근하는 두가지 방법인 GUI클라이언트, Python클라이언트에 대해 알아보았습니다