

본 자료와 관련 영상 콘텐츠는 저작권법 제25조 2항에 의해 보호를 받습니다.
본 콘텐츠 및 콘텐츠 일부 문구 등을 외부에 공개하거나, 요약해서 게시하지 말아주세요.
Copyright [잔재미코딩](#) Dave Lee

world database 로 익히는 데이터 분석과 SQL

- 다운로드: <https://dev.mysql.com/doc/index-other.html> 에서 world database ZIP 파일로 다운받아서, 압축 해제

world_x database 와 헷갈리지 않도록 유의해주세요~

world database sql 파일 실행하기

- Mysql Workbench
 - File -> Open SQL Script -> world.sql 선택 후 실행

world 테이블 정보 참고

- <https://downloads.mysql.com/docs/world-setup-en.a4.pdf>

실습 환경 맞추기

- Mysql Workbench 에서 Schemes 탭 선택 후, refresh 버튼을 눌러 world 데이터베이스 있는지 확인해보기

실습

테이블 정보를 확인해보세요

```
USE world
```

```
SHOW TABLES
```

```
SELECT * FROM city
```

전체 world 데이터베이스 테이블 확인해보기

실제 현업에서 사용하는 데이터베이스에는 다양한 테이블이 있지만, 각 테이블에 대한 매우 친절하고, 상세한 설명이 없는 경우도 많습니다. 데이터를 보며 유추하기도 합니다.

- 전체 데이터에 대한 꼼꼼한 설명은 <https://downloads.mysql.com/docs/world-setup-en.a4.pdf> 에 있음

SQL SELECT 문법1 : LIMIT

- 결과중 일부만 데이터 가져오기
 - 예) 결과중 처음부터 10개만 가져오기
 - SELECT * FROM 테이블이름 LIMIT 10 (테이블 데이터 중 최상위에 있는 10개의 데이터만 가져오기)
 - SELECT * FROM 테이블이름 WHERE 조건문 LIMIT 1 (특정 조건에 맞는 데이터 중 최상위에 있는 1개의 데이터만 가져오기)

연습문제를 풀기 위해 이해가 필요한 정보

- city 테이블 : 전세계 도시 정보를 포함하는 테이블
 - 테이블의 각 컬럼은 실제 컬럼값과 함께 확인해야 이해가 쉬움
 - 보통 다양한 컬럼이 있으므로, 이 중에 필요한 컬럼만 이해하면 됨

특정 테이블의 컬럼과 컬럼값 확인을 위해 LIMIT 1 을 많이 사용함

```
SELECT * FROM city LIMIT 1
```

SQL SELECT 문법2: COUNT

- 결과 수 세기 (데이터 행의 수 세기)
 - 예) 결과 수 세기
 - SELECT COUNT(*) FROM 테이블이름 (테이블 전체 데이터 수 세기)
 - SELECT COUNT(*) FROM 테이블이름 WHERE 조건문 (특정 조건에 맞는 테이블 데이터 수 세기)

연습문제1 (직접 해보신 후에, 영상으로 확인하세요!)

city 테이블에서 총 도시 수 알아내기

SQL SELECT 문법3: DISTINCT

- 특정 컬럼값 출력시 중복된 값을 출력하지 않음
 - 예) 유일한 컬럼값 확인하기
 - SELECT DISTINCT 컬럼명 FROM 테이블이름 (특정 컬럼에 들어가 있는 컬럼값 종류 확인하기)
 - SELECT DISTINCT 컬럼명 FROM 테이블이름 WHERE 조건문 (특정 조건에 맞는, 특정 컬럼에 들어가 있는 컬럼값 종류 확인하기)

연습문제2 (직접 해보신 후에, 영상으로 확인하세요!)

- city 테이블에서 국가코드 수 알아내기

연습문제3 (직접 해보신 후에, 영상으로 확인하세요!)

city 테이블에서 CountryCode 가 'KOR' 인 도시 이름 알아내기

연습문제4 (직접 해보신 후에, 영상으로 확인하세요!)

city 테이블에서 CountryCode 가 'KOR' 인 도시 수 알아내기

연습문제5 (직접 해보신 후에, 영상으로 확인하세요!)

city 테이블에서 Population 이 1000000 이상인 도시 수 알아내기

테이블과 테이블 간의 관계

- city, country, countrylanguage 와 같이 다른 테이블에 상세 정보가 있고, 각 정보를 구별할 수 있는 Code 값으로 연결되어 있는 경우가 많음

하나의 테이블에 모든 정보를 담지 않습니다.

```
SELECT * FROM city LIMIT 1
SELECT * FROM country LIMIT 1
SELECT * FROM countrylanguage LIMIT 1
```

연습문제6 (직접 해보신 후에, 영상으로 확인하세요!)

countrylanguage 테이블에서 전체 언어 수 알아내기

SQL SELECT 문법4: SUM, AVG, MAX, MIN

- 특정 컬럼값을 기반으로
 - SUM(): 컬럼값의 합계
 - AVG(): 컬럼값의 평균
 - MAX(): 컬럼값의 최대값
 - MIN(): 컬럼값의 최소값
 - 예) 컬럼값 분석하기
 - SELECT SUM(컬럼명) FROM 테이블이름 (특정 컬럼에 들어가 있는 컬럼값의 합계 구하기)

- SELECT AVG(컬럼명) FROM 테이블이름 WHERE 조건문 (특정 조건에 맞는, 특정 컬럼에 들어가 있는 컬럼값의 평균 구하기)

연습문제7 (직접 해보신 후에, 영상으로 확인하세요!)

- city 테이블에서 총 Population 알아내기
- city 테이블에서 최대 Population 알아내기
- city 테이블에서 최소 Population 알아내기
- city 테이블에서 평균 Population 알아내기

연습문제8 (직접 해보신 후에, 영상으로 확인하세요!)

city 테이블에서 CountryCode 가 'KOR' 인 도시의 총 Population 수 알아내기

SQL SELECT 문법5: GROUP BY

- 특정 컬럼값을 기반으로 그룹핑하기
 - 보통 SUM, COUNT, MAX, MIN 등과 함께 많이 사용됨
 - 예) SELECT SUM(Population) FROM city GROUP BY CountryCode (국가 코드로 그룹핑하면, 각 국가별 Population 수를 알아낼 수 있음)
 - 예) SELECT COUNT(*) FROM city GROUP BY CountryCode (국가 코드로 그룹핑하면, 국가 수를 알아낼 수 있음)

연습문제9 (직접 해보신 후에, 영상으로 확인하세요!)

각 국가별 도시 중 최대 Population 을 출력하시요

연습문제10 (직접 해보신 후에, 영상으로 확인하세요!)

각 국가별 총 Population 과 해당 국가코드를 출력하시요

연습문제11 (직접 해보신 후에, 영상으로 확인하세요!)

각 Continent 별 총 SurfaceArea 과 해당 Continent 를 출력하시요 (country 테이블 기반)

연습문제12 (직접 해보신 후에, 영상으로 확인하세요!)

각 Region 별 평균 GNP 와 해당 Region 를 출력하시요 (country 테이블 기반)

연습문제13 (직접 해보신 후에, 영상으로 확인하세요!)

각 Region 별 평균 LifeExpectancy 와 해당 Region 를 출력하시요 (country 테이블 기반)

연습문제14 (직접 해보신 후에, 영상으로 확인하세요!)

각 Region 별 최소 IndepYear 와 해당 Region 를 출력하시요 (country 테이블 기반)

SQL SELECT 문법6: ORDER BY

- 특정 컬럼값을 기준으로 데이터 정렬하기
 - ORDER BY 정렬할 기준 컬럼명 DESC|ASC
 - 예) SELECT * FROM city ORDER BY Population DESC (Population 값을 기준으로 내림차순으로 정렬해서 데이터를 가져와라)
 - 예) SELECT * FROM city ORDER BY Population ASC (Population 값을 기준으로 올림차순으로 정렬해서 데이터를 가져와라)
 - 예) SELECT * FROM city WHERE 조건문 ORDER BY Population ASC (특정 조건에 맞는 데이터를 Population 값을 기준으로 올림차순으로 정렬해서 데이터를 가져와라)
 - 예) SELECT COUNT(*) FROM city WHERE 조건문 GROUP BY 컬럼 ORDER BY COUNT(*) ASC (특정 조건에 맞는 데이터를 특정 컬럼값을 기준으로 그룹핑하되, Population 값을 기준으로 올림차순으로 정렬해서 데이터를 가져와라)
 - DESC 또는 ASC 를 안쓰면 ASC 로 정렬

SQL 조건 순서

```
SELECT 컬럼
FROM 테이블명
WHERE 조건
GROUP BY 컬럼
ORDER BY 컬럼
LIMIT
```

연습문제15 (직접 해보신 후에, 영상으로 확인하세요!)

country 테이블에서 Region 별로, IndepYear 가 1900 이상인 국가만 대상으로 평균 GNP 이 낮은 순으로 5개의 Region 만 Region 과 평균 GNP 를 출력하세요 (이 문제를 풀기 위해, 연습문제16 ~ 19까지 작성해보신 후에, 한번에 연습문제 15번을 작성해보세요)

연습문제16 (직접 해보신 후에, 영상으로 확인하세요!)

IndepYear 가 1900 이상인 국가를 출력하세요

연습문제17 (직접 해보신 후에, 영상으로 확인하세요!)

IndepYear 가 1900 이상인 국가를 Region 별로 묶어서, 평균 GNP 를 출력하세요

연습문제18 (직접 해보신 후에, 영상으로 확인하세요!)

IndepYear 가 1900 이상인 국가를 Region 별로 묶어서, 평균 GNP 이 낮은 순으로 평균 GNP를 출력하세요

연습문제19 (직접 해보신 후에, 영상으로 확인하세요!)

IndepYear 가 1900 이상인 국가를 Region 별로 묶어서, 평균 GNP 이 낮은 순으로 5개의 Region 만 평균 GNP를 출력

하세요

SQL SELECT 문법7: AS

- 표시할 컬럼명도 다르게 하기
 - 예) SELECT COUNT(*) AS total FROM city (city의 전체 데이터 갯수를 COUNT(*) 로 표시하지 말고, total 로 표시하기)

연습문제20 (직접 해보신 후에, 영상으로 확인하세요!)

Region 테이블에서 Region 별로, IndepYear 가 1900 이상인 국가만 대상으로 평균 GNP 이 낮은 순으로 5개의 Region 만 Region 과 평균 GNP 를 출력하세요. 단 평균 GNP 컬럼명을 AvgGNP 로 출력하세요

본 자료와 관련 영상 콘텐츠는 저작권법 제25조 2항에 의해 보호를 받습니다.

본 콘텐츠 및 콘텐츠 일부 문구 등을 외부에 공개하거나, 요약해서 게시하지 말아주세요.

Copyright [잔재미코딩](#) Dave Lee