

R 프로그래밍 중급으로 나아가기 위한 Level UP Skill

Step2. 하나의 문법으로 다양한 소스의 데이터를 다뤄보기

https://mrchypark.github.io/dabrp_classnote3/class2

박찬엽

2017년 9월 21일

목차

1. 과제 질답
2. 수업의 목표
3. 데이터 공유와 탐색
 - 데이터를 공유하는 방법
 - 실습에 사용할 데이터 소개
 - 실습
4. 데이터 베이스와 R의 연결
 - DBI 소개와 SQLite
 - DBI 함수의 동작
 - 실습
5. 본격 DB와 R의 연결
 - 클라우드 서비스 이용하기
 - MySQL 연결
 - BigQuery 연결
6. 과제

과제 질답

수업의 목표

1. 데이터를 잘 공유하는 법을 알고 설명할 수 있다.
2. DBI 패키지의 구조를 이해하고 설명할 수 있다.
3. 클라우드 서비스에 테이블을 저장하고 R로 연동할 수 있다.

데이터를 공유하기

데이터 공유

Leek group에서 소개하는 데이터 공유 가이드

- 원시 데이터(raw data) : 최초에 수집된 데이터를 뜻하며 센서나 인터넷 서비스의 로그 데이터나 사용자가 직접 작성한 설문 데이터 등을 의미함
- 정제후 데이터(tidy data) : 사용하기 좋은 상태로 정리한 데이터를 뜻하며 보통 설문조사에서는 코딩한다고 표현함
- 코드북(code book) : 공개하는 데이터에 대해서 설명하는 문서로 단위나 각 변수명이 의미하는 바 등을 작성함
- 자료형(data type) : 변수의 종류와 저장법에 대해서 조언하며 엑셀등 특별한 프로그램에 종속적인 기능은 사용하지 말 것을 권고함
- 재현성(reproducibility) : 다시 결과를 만들 수 있는 것을 뜻하며 2010년대 이후 과학적 방법의 중요한 축

원시 데이터

최초 획득한 당시 그대로의 데이터

- 어떤 식으로든 수정을 하지 않은 상태
- 수정을 하는 과정을 함께 기록함으로써 신뢰성 확보
- 위 두 가지가 없는 경우 상황을 상상해야 함

정제후 데이터

해들리 위컴이 설명한 **tidy data**의 요건에 맞게 가공하여 데이터를 쉽게 다룰 수 있게 만든 상태

- 측정하는 각 변수는 하나의 열에 있어야 함
- 측정하는 각 관찰은 하나의 행에 있어야 함
- 각 종류의 변수에 대해 각 하나의 테이블이 있어야 함
- 여러 개의 테이블이 있는 경우 테이블에 합치기 위한 기준 열을 포함해야 함

코드북

데이터셋에 대해 필요한 설명을 담은 문서

- 정제후 데이터에 대해 추가적으로 필요한 설명이나 정보(단위 등)
- 정제 과정에서 사용한 방법의 설명과 사용한 이유
- 데이터가 사용된 분석에 대한 정보

데이터 원본

확보할 당시의 원시 데이터나, 항상 최신 상태를 유지하여 신뢰할 수 있는 데이터. 기계적 생성 데이터가 아닌 사람이 원본이 되므로 작성된 내용이 원본이라는 합의를 함으로써 데이터 관리의 체계를 세울 수 있음

- 커뮤니케이션 비용 감소
- 의사결정 및 활동의 기준
- 가공된 데이터의 신뢰성 확보

데이터 훑어보기

- `head()`: 최초 6행의 데이터를 보여줌(행갯수 조절 가능)
- `tail()`: 마지막 6행의 데이터를 보여줌(행갯수 조절 가능)
- `summary()`: 각 컬럼의 자료형과 숫자라면 대표값을 함께 보여줌
- `str()`: 각 컬럼의 자료형과 초기 값을 보여줌
- `length()`: 데이터의 길이 출력(vector)
- `nrow()`: 행 갯수 출력(data.frame)
- `is.na()`: NA 인지 확인
- `complete.cases()`: 값이 모두 있는지 행단위로 검사
- `tibble()`: 최근 기법으로 재구성한 tidyverse 패키지 군에서 사용하는 data.frame의 일종
- `data.table()`: DB에서 사용하는 기법 등과 메모리 효율성을 극대화한 data.frame의 일종

데이터 소개

실습 데이터 설정하기

```
# data.table 패키지가 있는지 확인하고 없으면 설치합니다.  
if (!requireNamespace("data.table")) install.packages("data.table")  
# data.table 패키지를 사용할 수 있게 불러옵니다.  
library(data.table)  
  
# tran.csv 파일이 data/recomen 폴더에 있는지 확인 합니다.  
chk<-dir("../data/recomen", pattern = "tran.csv")  
# 없으면 다운로드합니다.  
if(identical(chk, character(0))){  
  recoment<-"http://rcoholic.duckdns.org/oc/index.php/s/jISrPutj4oCLci2/download"  
  download.file(recoment, destfile="../data/recomen/tran.csv", mode='wb')  
}  
# fread 함수를 이용해서 csv 파일을 R 객체로 불러옵니다.  
chennel<-fread("../data/recomen/chennel.csv")  
competitor<-fread("../data/recomen/competitor.csv")  
customer<-fread("../data/recomen/customer.csv")  
item<-fread("../data/recomen/item.csv")  
membership<-fread("../data/recomen/membership.csv")  
tran<-fread("../data/recomen/tran.csv")
```

추천 데이터 - chennel

- cusID : 5자리 숫자조합으로 구성된 고객ID
- chennel: 접속 체널
- useCnt : 사용횟수(건)

```
str(chennel, max.level=1)
```

```
## Classes 'data.table' and 'data.frame': 8824 obs. of 3 variables:  
## $ cusID : int 7 14 42 74 74 94 112 112 122 123 ...  
## $ chennel: chr "A_MOBILE/APP" "A_MOBILE/APP" "B_MOBILE/APP" "A_MOBILE/APP" ...  
## $ useCnt : int 4 1 23 1 30 14 16 1 27 10 ...  
## - attr(*, ".internal.selfref")=<externalptr>
```

추천 데이터 - competitor

- cusID : 5자리 숫자조합으로 구성된 고객ID
- partner : 제휴사
- competitor: 경쟁사
- useDate : 이용년월(YYYYDD)

```
str(competitor, max.level=1)
```

```
## Classes 'data.table' and 'data.frame': 28159 obs. of 4 variables:
## $ cusID     : int  2 51 77 77 77 77 77 88 88 110 ...
## $ partner   : chr  "D" "D" "D" "D" ...
## $ competitor: chr  "D02" "D01" "D02" "D02" ...
## $ useDate   : int  201507 201504 201503 201506 201507 201508 201511 201510 201511 201508 ...
## - attr(*, ".internal.selfref")=<externalptr>
```

추천 데이터 - customer

- cusID: 5자리 숫자조합으로 구성된 고객ID
 - sex : 성별
 - age : 연령
 - area : 거주지역

```
str(customer, max.level=1)
```

추천 데이터 - item

- partner : 재휴사
- cate_1 : 대분류
- cate_2 : 중분류
- cate_3 : 소분류
- cate_2_name: 중분류명
- cate_3_name: 소분류명

```
str(item, max.level=1)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':      4386 obs. of  6 variables:
##   $ partner     : chr  "A" "A" "A" "A" ...
##   $ cate_1      : chr  "01" "01" "01" "01" ...
##   $ cate_2      : chr  "0101" "0101" "0101" "0101" ...
##   $ cate_3      : chr  "A010101" "A010102" "A010103" "A010104" ...
##   $ cate_2_name: chr  "일용잡화" "일용잡화" "일용잡화" "일용잡화" ...
##   $ cate_3_name: chr  "위생세제" "휴지류" "뷰티상품" "일용잡화" ...  
## - attr(*, "spec")=List of 2  
##   ..- attr(*, "class")= chr "col_spec"
```

추천 데이터 - membership

- cusID : 5자리 숫자조합으로 구성된 고객ID
- memberShip: 멤버십명
- regDate : 가입년월

```
str(membership, max.level=1)
```

```
## Classes 'data.table' and 'data.frame': 7456 obs. of 3 variables:  
## $ cusID      : int 11 21 37 43 44 61 65 69 76 79 ...  
## $ memberShip: chr "\xed\\uc5d0\ud55c\uc11c" "\xed\uc5d0\ud55c\uc11c" "\xed\uc5d0\ud55c\uc11c" "\xed\uc5d0\ud55c\uc11c" ...  
## $ regDate    : int 201512 201506 201306 201403 201411 201312 201506 201404 201406 201311 ...  
## - attr(*, ".internal.selfref")=<externalptr>
```

추천 데이터 - tran

```
summary(tran)
```

```
##   partner           receiptNum        cate_1        cate_2
## Length:28593030    Min.   : 1      Min.   : 1.00  Min.   : 101
## Class  :character  1st Qu.: 3922474 1st Qu.: 4.00  1st Qu.: 401
## Mode   :character  Median : 7167787 Median :11.00  Median :1102
##                   Mean   : 6447881 Mean   :18.37  Mean   :1840
##                   3rd Qu.: 9116336 3rd Qu.:18.00  3rd Qu.:1808
##                   Max.  :11096601 Max.  :92.00  Max.  :9206
##   cate_3            cusID       storeCode       date
## Length:28593030    Min.   : 1      Min.   : 1.00  Min.   :20140101
## Class  :character  1st Qu.: 5206   1st Qu.: 16.00 1st Qu.:20140711
## Mode   :character  Median :10104   Median : 44.00  Median :20150110
##                   Mean   : 9904   Mean   : 92.26  Mean   :20145817
##                   3rd Qu.:14638   3rd Qu.:110.00 3rd Qu.:20150703
##                   Max.  :19383   Max.  :593.00  Max.  :20151231
##   time              amount
## Min.   : 0.00      Min.   : 1
## 1st Qu.:14.00      1st Qu.: 2050
## Median :17.00      Median : 4290
## Mean   :16.71      Mean   : 23678
## 3rd Qu.:19.00      3rd Qu.: 9900
## Max.  :23.00      Max.  :101330000
```

추천 데이터 - tran

```
str(tran, max.level=1)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame': 28593030 obs. of 10 variables:
## $ partner : chr "B" "B" "B" "B" ...
## $ receiptNum: int 8664000 8664000 8664000 8664000 8664001 8664001 8664002 8664002 8664002 8664003
## $ cate_1   : int 15 16 16 18 5 15 10 43 54 5 ...
## $ cate_2   : int 1504 1601 1602 1803 509 1501 1003 4301 5403 504 ...
## $ cate_3   : chr "B150401" "B160101" "B160201" "B180301" ...
## $ cusID    : int 17218 17218 17218 17218 17674 17674 14388 14388 14388 15773 ...
## $ storeCode: int 44 44 44 44 44 44 44 44 44 ...
## $ date     : int 20140222 20140222 20140222 20140222 20140222 20140222 20140222 20140222 20140222 20140222
## $ time     : int 20 20 20 20 22 22 23 23 23 21 ...
## $ amount   : int 2420 1070 8060 6000 1120 1200 5290 5960 9900 970 ...
## - attr(*, "spec")=List of 2
##   ..- attr(*, "class")= chr "col_spec"
```

실습

1. dabrp_classnote3 프로젝트를 pull 하여 data 폴더를 받고 확인하세요
2. 13p의 실습데이터 설정 스크립트를 실행하여 tran.csv(1.4G) 파일을 받으세요
3. 11p의 함수들을 활용하여 데이터를 탐색해보세요
* tibble이나 data.table은 다음 시간에 다룰 예정입니다

데이터베이스

데이터란

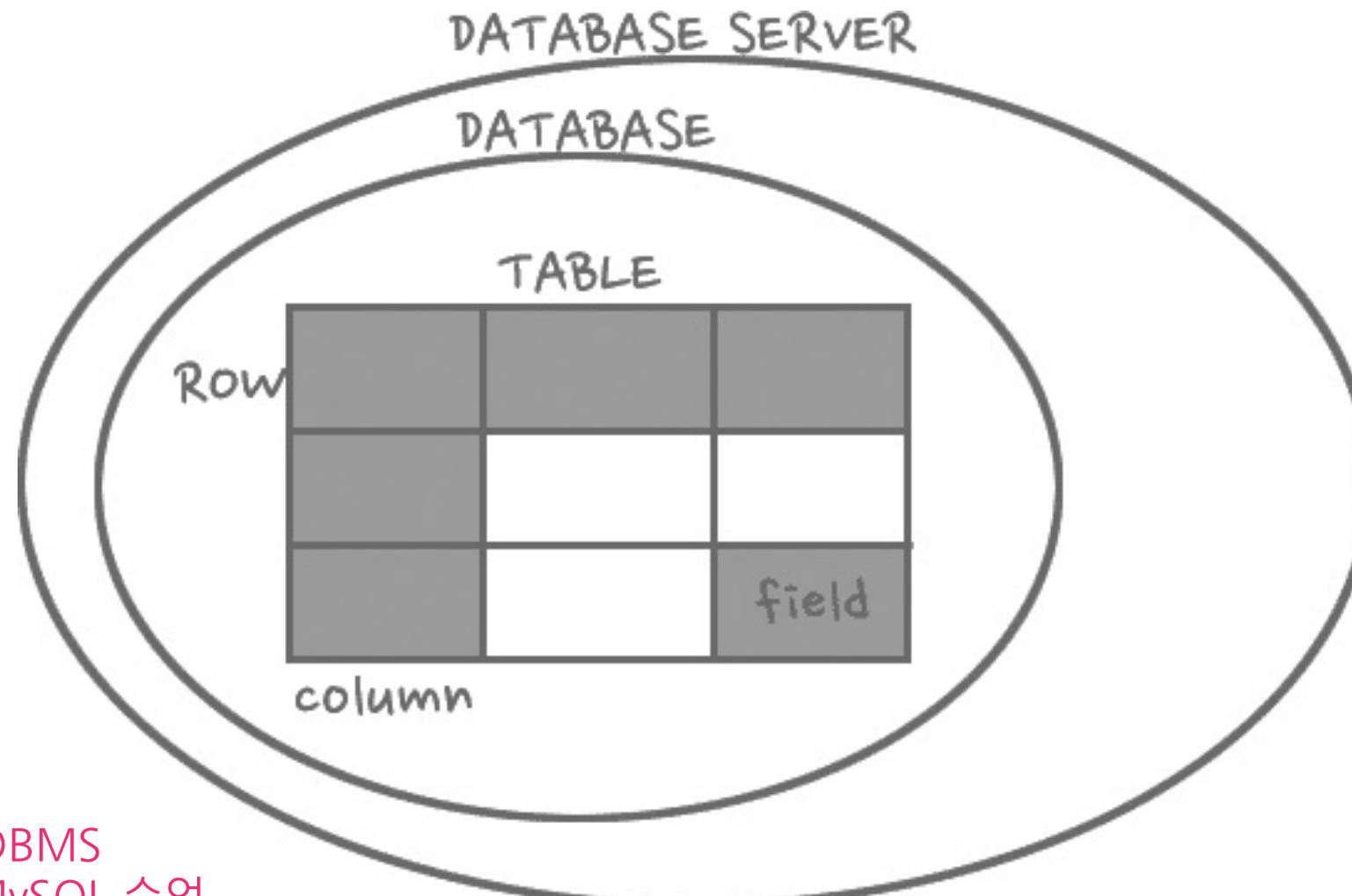
단순한 관찰이나 측정 등의 수단을 통해 현실 세계로부터 수집된 사실이나 값. 의미있게 사용하기 위해서 구조화*가 필요함



* 구조화: 체계적으로 조직하는 것

Database Management System(DBMS)¹

DBMS는 데이터베이스를 관리하는 응용프로그램의 일종으로 중복 제어, 접근 권한 관리, 무결성 제약, 데이터 관리 등의 기능을 제공하며 제품마다 특성이 매우 다양함

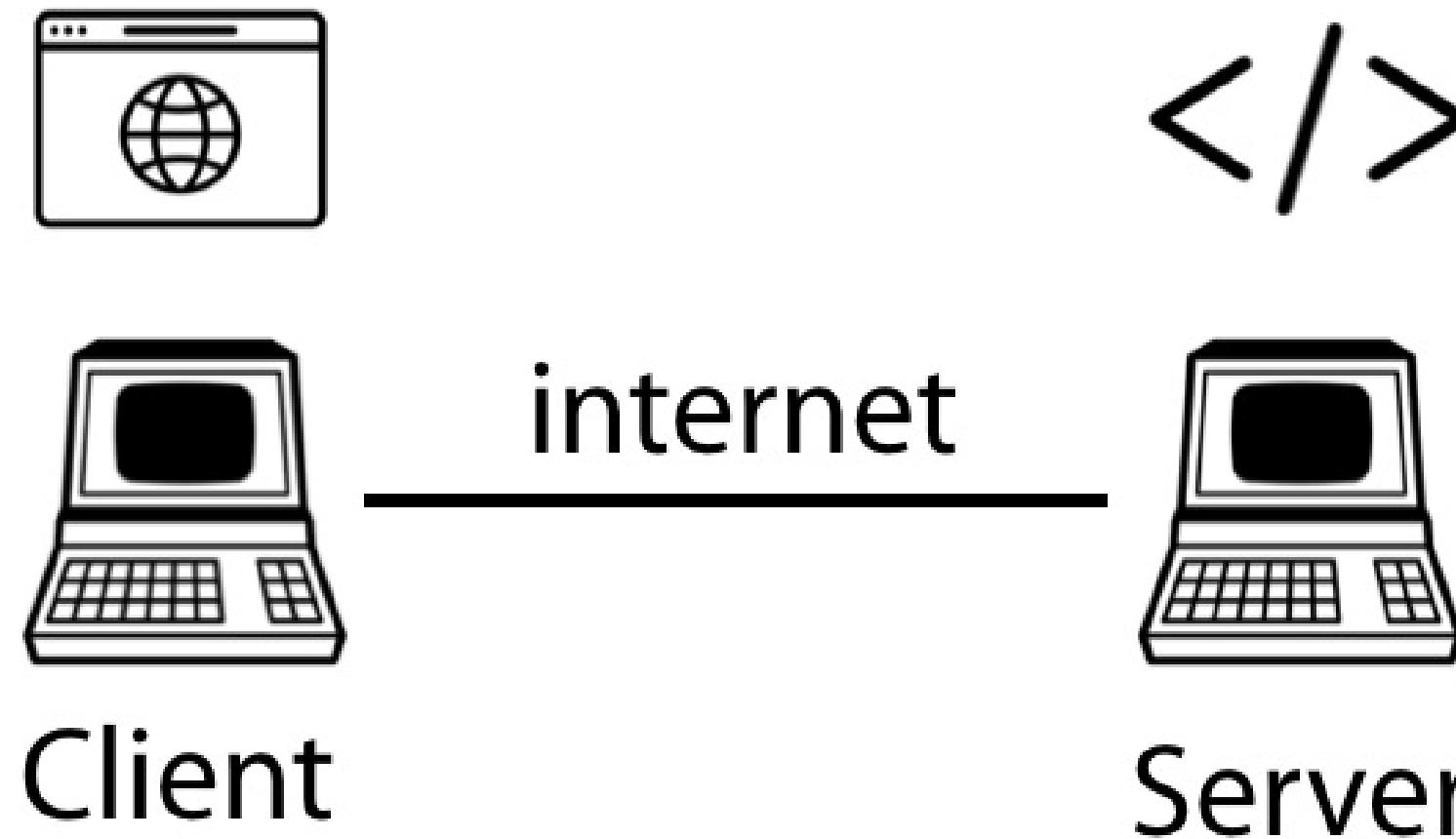


1 <https://namu.wiki/w/DBMS>

* 이미지 출처: 생활코딩 MySQL 수업

서버와 클라이언트

유명하게 사용되는 DBMS는 잘 활용되는 GUI가 있기도 하나 그 기능 구현이 서버와 클라이언트로 되어 있어 연결시 ip 주소가 필요하는 등 관련 지식이 필요함



데이터베이스 클라이언트

- 대표적인 클라이언트
 - MySQL monitor
 - PHPmyAdmin
 - Navicat
 - HeidiSQL

테이블(table)

- 2차원의 자료형으로 table과 같이 행(row)과 열(column)이 있는 엑셀의 시트와 비슷한 모양

```
print(head(iris))
```

```
##   Sepal.Length Sepal.Width  Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4        0.2  setosa
## 2          4.9         3.0          1.4        0.2  setosa
## 3          4.7         3.2          1.3        0.2  setosa
## 4          4.6         3.1          1.5        0.2  setosa
## 5          5.0         3.6          1.4        0.2  setosa
## 6          5.4         3.9          1.7        0.4  setosa
```

* R 데이터 객체인 iris의 data.frame 예시

DBI(database interface)

DBI는 odbc나 java의 jdbc, 파이썬의 db-api와 같이 R에서 데이터베이스를 컨트롤하기 위한 인터페이스 패키지

```
if (!require(devtools)) install.packages("devtools")

## Loading required package: devtools

if (!require(DBI)) devtools::install_github("rstats-db/DBI")

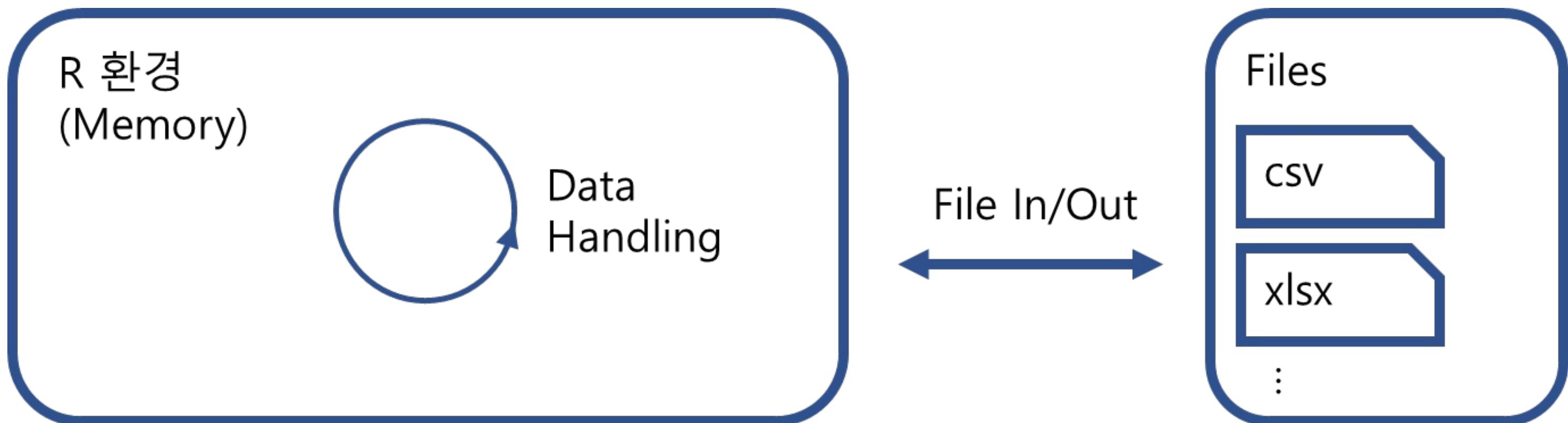
## Loading required package: DBI

if (!require(RSQLite)) devtools::install_github("rstats-db/RSQLite")

## Loading required package: RSQLite
```

파일의 입출력

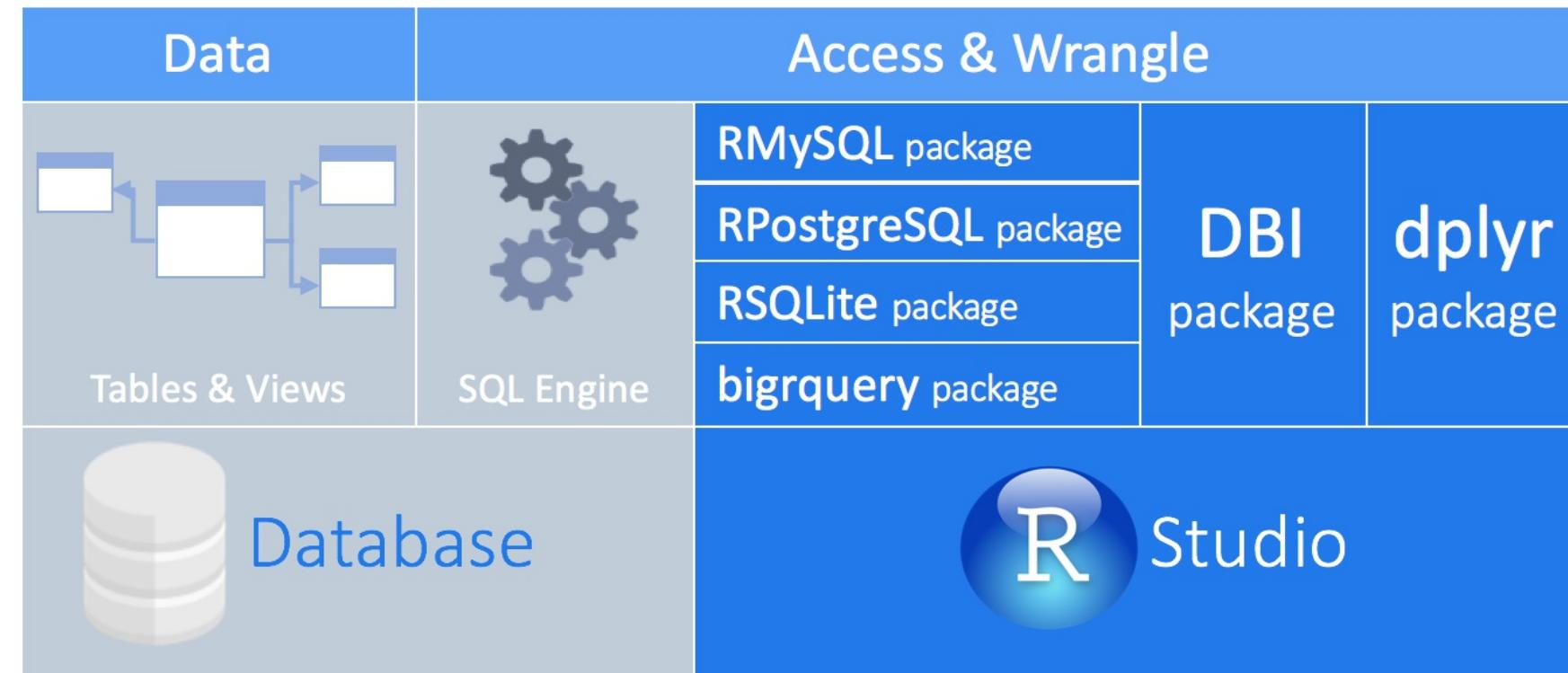
R은 데이터를 사용하기 위해 메모리로 구성된 R 환경에 파일의 내용을 R객체 형태의 데이터로 변환해서 사용



DBI의 DBMS 연결

DBI(Database Interface) 패키지는 각 DBMS에 해당하는 드라이버 패키지와 연동하여 DBMS를 통해 데이터를 사용

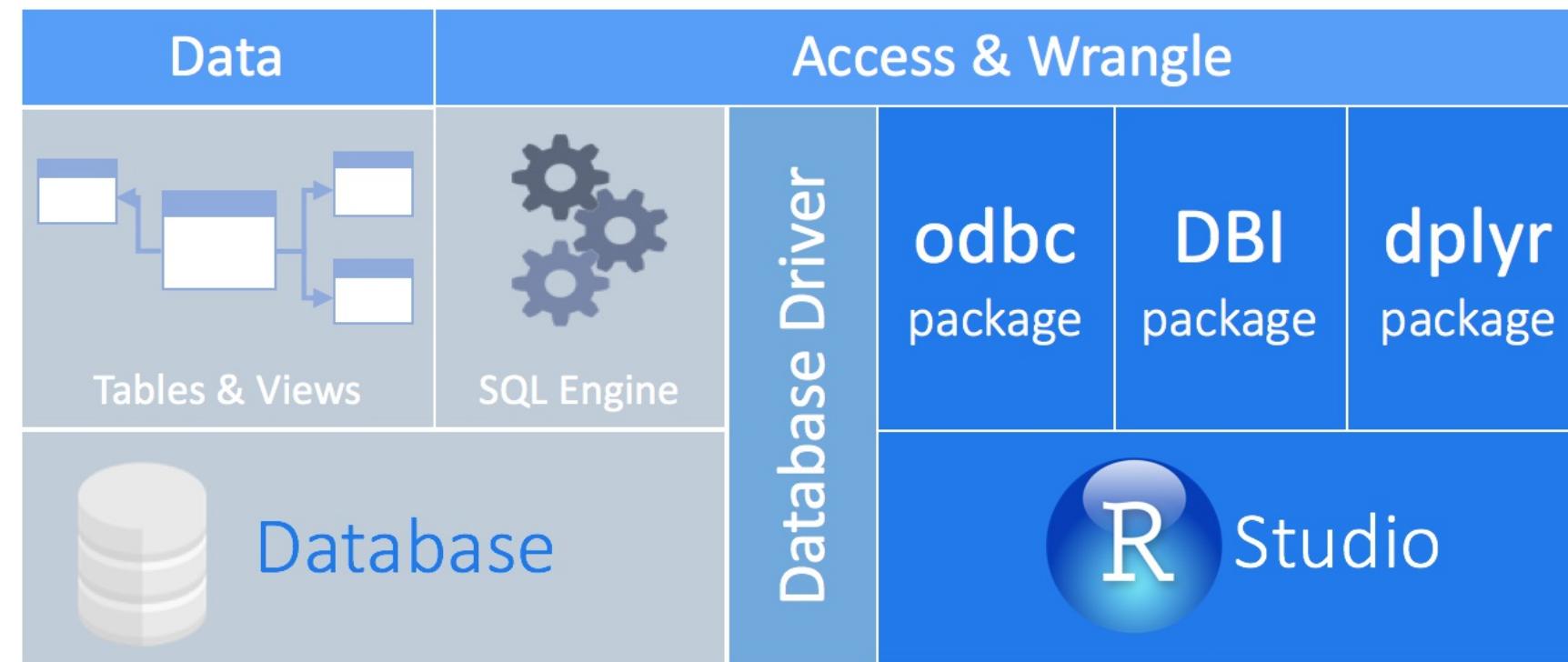
Open Source Databases



ODBC를 통한 연결

ODBC를 직접 지원하기도 함

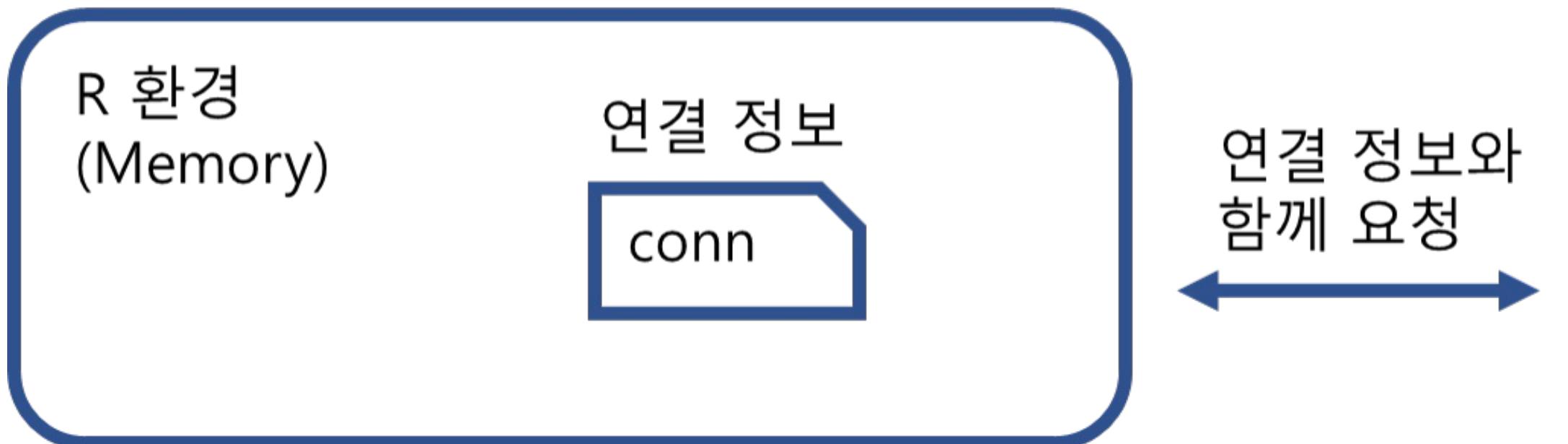
Commercial Databases



데이터 베이스와 연결

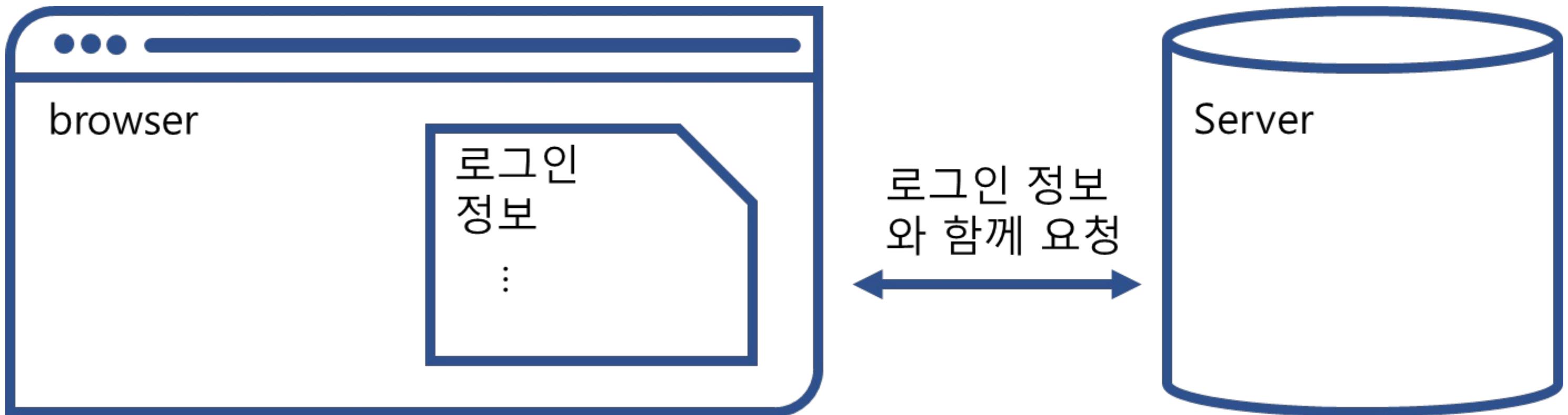
파일과 다르게 데이터베이스는 DBMS를 통해서 내부의 데이터에 접근해야 하기 때문에 DBMS의 연결 정보를 가진 conn 객체를 만들고 그것을 활용하는 명령 구조를 따름

```
conn <- dbConnect( drv, ... )
```



데이터 베이스와 연결

정확히 같진 않지만 `conn`은 브라우저가 로그인 정보등을 바탕으로 서버와 통신하는 것과 비슷하게 동작함



SQLite

SQLite is a self-contained, high-reliability, embedded, full-featured, public-domain, SQL database engine.

파일 하나로 구성하는 작고, 무료인 sql db

DBI 함수

패키지 불러오기. 패키지가 설치되어 있지 않으면 28p 참조

```
library(DBI)
library(RSQLite)
library(data.table)
```

SQLite 와 연결하기

conn은 관례적으로 연결 정보를 저장하는 객체의 이름이며 SQLite는 DB가 파일로 구성되어 있어서 권한(ID나 비밀번호) 설정이 없고 연결 정보도 단순(host 없음)함

```
conn <- dbConnect(RSQLite::SQLite(), dbname = ".../ForderForClass2/class2.sqlite")
conn

## <SQLiteConnection>
##   Path: C:\Users\mrchypark\Documents\project\dabrp_classnote3\ForderForClass2\class2.sqlite
##   Extensions: TRUE

# DB내에 있는 테이블의 이름을 확인함
dbListTables(conn)

## character(0)
```

DBI 명령으로 table 만들기

DBI 패키지에는 테이블을 작성하고 지우는 등 원래 SQL로 작성하는 Query의 몇 가지 유용한 경우를 함수로 만들어 제공.
`dbwriteTable()` 함수는 R 객체를 바로 테이블로 작성하는 기능

```
dbwriteTable(conn, "mtcars", mtcars, overwrite = T)  
dbListTables(conn)  
## [1] "mtcars"
```

DBI 명령으로 table 불러오기

dbReadTable()은 명령어에서도 이해하기 쉽듯 DB내에 있는 특정 테이블을 이름으로 지정하여 데이터를 불러오는 함수

```
dbReadTable(conn, "mtcars")
```

```
##   mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## 1 21.0   6 160.0 110 3.90 2.620 16.46  0  1     4    4
## 2 21.0   6 160.0 110 3.90 2.875 17.02  0  1     4    4
## 3 22.8   4 108.0  93 3.85 2.320 18.61  1  1     4    1
## 4 21.4   6 258.0 110 3.08 3.215 19.44  1  0     3    1
## 5 18.7   8 360.0 175 3.15 3.440 17.02  0  0     3    2
## 6 18.1   6 225.0 105 2.76 3.460 20.22  1  0     3    1
## 7 14.3   8 360.0 245 3.21 3.570 15.84  0  0     3    4
## 8 24.4   4 146.7  62 3.69 3.190 20.00  1  0     4    2
## 9 22.8   4 140.8  95 3.92 3.150 22.90  1  0     4    2
## 10 19.2   6 167.6 123 3.92 3.440 18.30  1  0     4    4
## 11 17.8   6 167.6 123 3.92 3.440 18.90  1  0     4    4
## 12 16.4   8 275.8 180 3.07 4.070 17.40  0  0     3    3
## 13 17.3   8 275.8 180 3.07 3.730 17.60  0  0     3    3
## 14 15.2   8 275.8 180 3.07 3.780 18.00  0  0     3    3
## 15 10.4   8 472.0 205 2.93 5.250 17.98  0  0     3    4
## 16 10.4   8 460.0 215 3.00 5.424 17.82  0  0     3    4
## 17 14.7   8 440.0 230 3.23 5.345 17.42  0  0     3    4
## 18 32.4   4  78.7  66 4.08 2.200 19.47  1  1     4    1
## 19 30.4   4  75.7  52 4.93 1.615 18.52  1  1     4    2
## 20 33.9   4  71.1  65 4.22 1.835 19.90  1  1     4    1
## 21 21.5   4 120.1  97 3.70 2.465 20.01  1  0     3    1
## 22 15.5   8 318.0 150 2.76 3.520 16.87  0  0     3    2
## 23 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 24 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 25 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 26 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 27 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 28 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 29 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 30 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 31 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 32 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 33 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 34 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 35 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 36 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 37 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 38 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 39 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 40 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 41 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 42 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 43 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 44 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 45 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 46 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 47 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 48 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 49 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 50 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 51 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 52 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 53 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 54 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 55 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 56 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 57 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 58 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 59 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 60 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 61 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 62 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 63 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 64 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 65 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 66 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 67 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 68 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 69 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 70 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 71 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 72 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 73 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 74 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 75 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 76 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 77 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 78 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 79 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 80 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 81 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 82 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 83 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 84 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 85 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 86 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 87 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 88 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 89 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 90 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 91 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 92 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 93 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 94 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 95 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 96 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 97 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 98 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 99 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
## 100 15.2   8 304.0 150 3.15 3.435 17.30  0  0     3    2
```

DBI 명령으로 table 지우기

`dbRemoveTable()` 은 테이블을 이름으로 지정하여 지우는 함수

```
dbRemoveTable(conn, "mtcars")
```

```
dbListTables(conn)
```

```
## character(0)
```

실습

1. DBI 패키지로 연결, 쓰기, 읽기를 수행해 봅니다.

1. **dbConnect()**함수와 **SQLite()**함수를 이용해서 "./ForderForclass2" 폴더에 "db_[자기이름영어].sqlite" 파일 이름으로 "conn" 연결 정보를 만드세요.
2. "dbListTables(conn)"으로 테이블이 아무것도 없는 것을 확인하세요.
3. "./data/recomen/chennel.csv"을 "chen" 이름의 R 객체로 불러오세요.
4. "dbWriteTable()" 함수를 이용해 "chen" 데이터 객체를 DB에 "dbchen"이라는 테이블 이름으로 저장해 보세요.
5. "dbReadTable()" 함수로 "dbchen"을 불러와서 "chen" 데이터 객체와 내용이 같은지 확인하세요.
* identical() 함수는 두 개의 객체가 같은지를 T/F로 결과를 주는 함수입니다.
6. "./ForderForclass2" 폴더에 "db_[자기이름영어].sqlite" 이 "chen" 테이블을 저장한 상태를 commit 해주세요.
7. github에서 PR을 진행해주세요.
8. 같은 과정을 "./data/recomen/customer.csv" 파일로 진행해 주세요. 중간에 명명은 "cust", "dbcust" 입니다.
* 인코딩 문제가 발생할 수 있습니다.

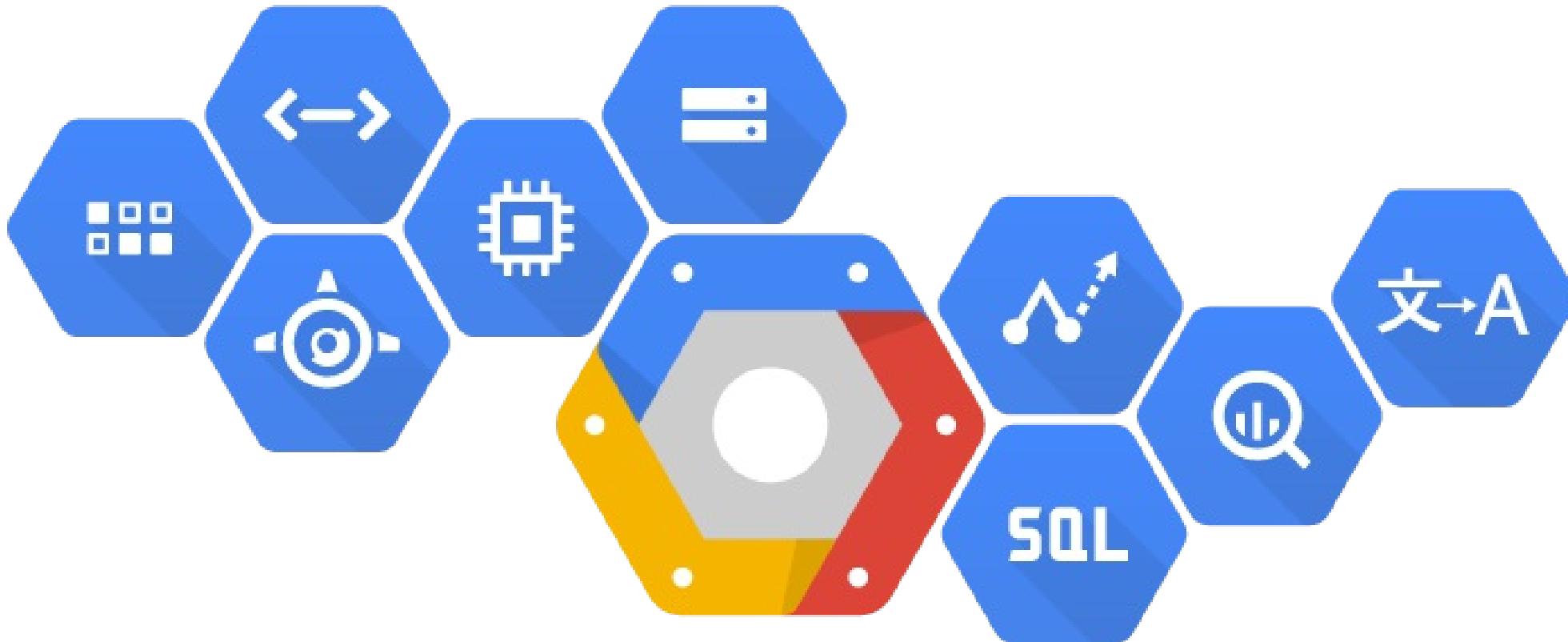
클라우드 서비스

클라우드 서비스 소개

클라우딩 컴퓨팅은 사용자의 환경 밖에서 서비스로서 제공된 확장 가능한 컴퓨팅 자원을 사용한 양에 따라 비용을 지불하고 사용하는 것

출처: P. Changanti, 가상 인프라용 클라우드 서비스, Part 1: IaaS(Infrastructure as a Service) 및 Eucalyptus

구글 클라우드



Google Cloud Platform

구글 클라우드 시작하기

<https://cloud.google.com/>

The screenshot shows the Google Cloud Platform homepage. At the top, there's a navigation bar with links for Why Google, Products, Solutions, Launcher, Pricing, Customers, Documentation, Support, and Partners. On the right of the navigation bar are 'TRY IT FREE' and 'CONTACT SALES' buttons. The main headline reads 'Build What's Next Better software. Faster.' Below this, there's a bulleted list of benefits: '✓ Use Google's core infrastructure, data analytics and machine learning.', '✓ Secure and fully featured for all enterprises.', and '✓ Committed to open source and industry leading price-performance.' At the bottom, there are three sections: 'GCP Region Expansion', 'Cloud Video Intelligence', and 'Machine Learning Webinar'. Each section has a brief description and a link.

Google Cloud Computin x

← → ⌂ 🔒 안전함 | https://cloud.google.com

Google Cloud Platform

Why Google Products Solutions Launcher Pricing Customers Documentation Support Partners TRY IT FREE CONTACT SALES

Build What's Next
Better software. Faster.

- ✓ Use Google's core infrastructure, data analytics and machine learning.
- ✓ Secure and fully featured for all enterprises.
- ✓ Committed to open source and industry leading price-performance.

TRY IT FREE CONTACT SALES

GCP Region Expansion

Google Cloud Platform customers can now run their workloads in more locations around the world.

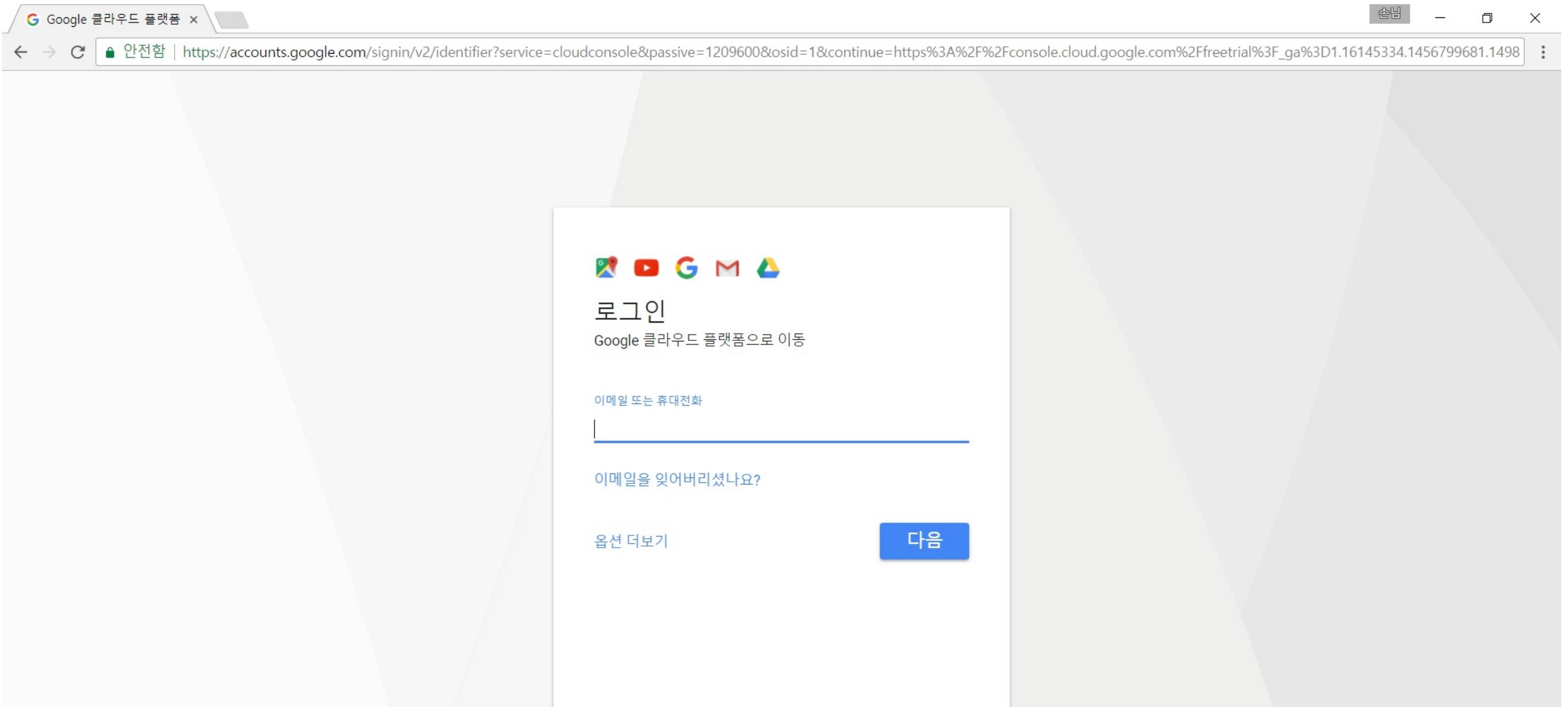
Cloud Video Intelligence

Understand the content of a video in an easy to use REST API.

Machine Learning Webinar

Join us live as Google and MIT discuss the value of ML and offer practical advice on getting started today.

구글 클라우드 로그인



The screenshot shows a web browser window for Google Cloud Platform. The title bar says "Google 클라우드 플랫폼". The address bar shows a secure connection to "https://accounts.google.com/signin/v2/identifier?service=cloudconsole&passive=1209600&osid=1&continue=https%3A%2F%2Fconsole.cloud.google.com%2Ffreetrial%3F_ga%3D1.16145334.1456799681.1498". The main content is a login form for Google Cloud Platform. It features several Google service icons at the top: Google Maps, YouTube, Google Search, Gmail, and Google Drive. Below them is the text "로그인" (Login) and "Google 클라우드 플랫폼으로 이동" (Move to Google Cloud Platform). A text input field is labeled "이메일 또는 휴대전화" (Email or phone number). Below the input field is a link "이메일을 잊어버리셨나요?" (Forgot your email?). At the bottom left is a link "옵션 더보기" (More options), and at the bottom right is a blue "다음" (Next) button. The footer of the page includes links for "한국어" (Korean), "도움말" (Help), "개인정보 보호" (Privacy), and "약관" (Terms).

Google 클라우드 플랫폼

안전함 | https://accounts.google.com/signin/v2/identifier?service=cloudconsole&passive=1209600&osid=1&continue=https%3A%2F%2Fconsole.cloud.google.com%2Ffreetrial%3F_ga%3D1.16145334.1456799681.1498

로그인

Google 클라우드 플랫폼으로 이동

이메일 또는 휴대전화

이메일을 잊어버리셨나요?

옵션 더보기

다음

한국어 ▾

도움말

개인정보 보호

약관

구글 클라우드 설문

The screenshot shows a web browser window for the Google Cloud Platform free trial. The URL in the address bar is https://console.cloud.google.com/freetrial?_ga=1.16145334.1456799681.1498616668&pli=1&page=0. The page title is "Google Cloud Platform".

Cloud Platform 무료로 사용해 보기

국가: 대한민국

수락: 기능 공지사항, 성능 제안사항, 의견 설문조사 및 특별 할인에 대한 새로운 소식을 이메일로 받고 싶습니다.

예 아니요

Google Cloud Platform 무료 평가판 서비스 약관을 읽었으며 이에 동의합니다.
계속 진행하려면 체크박스를 선택하세요.

예 아니요

동의 및 계속하기

모든 Cloud Platform 제품에 액세스
앱, 웹사이트, 서비스를 구축하고 실행하는 데 필요한 Firebase, Google Maps API 등의 모든 기능을 이용할 수 있습니다.

무료 크레딧 300달러
가입하여 Google Cloud Platform에서 이후 12개월 동안 사용할 수 있는 300달러의 혜택을 받아 보세요.

무료 평가판 종료 후 자동 청구되지 않음
신용카드를 요청하는 이유는 자동 가입을 방지하기 위해서입니다. 유료 계정으로 직접 업그레이드하지 않는 한 요금이 청구되지 않습니다.

개인정보처리방침

46 / 99

구글 클라우드 설문 선택

The screenshot shows a web browser window for the Google Cloud Platform free trial. The URL in the address bar is https://console.cloud.google.com/freetrial?_ga=1.16145334.1456799681.1498616668&pli=1&page=0. The page title is "Google Cloud Platform".

Cloud Platform 무료로 사용해 보기

Google

국가
내안민국

수락
기능 공지사항, 성능 제안사항, 의견 설문조사 및 특별 할인에 대한 새로운 소식을 이메일로 받고 싶습니다.
 예 아니요

Google Cloud Platform 무료 평가판 서비스 약관을 읽었으며 이에 동의합니다.
계속 진행하려면 체크박스를 선택하세요.
 예 아니요

동의 및 계속하기

모든 Cloud Platform 제품에 액세스
앱, 웹사이트, 서비스를 구축하고 실행하는 데 필요한 Firebase, Google Maps API 등의 모든 기능을 이용할 수 있습니다.

무료 크레딧 300달러
가입하여 Google Cloud Platform에서 이후 12개월 동안 사용할 수 있는 300달러의 혜택을 받아 보세요.

무료 평가판 종료 후 자동 청구되지 않음
신용카드를 요청하는 이유는 자동 가입을 방지하기 위해서입니다. 유료 계정으로 직접 업그레이드하지 않는 한 요금이 청구되지 않습니다.

개인정보처리방침

47 / 99

구글 클라우드 카드 등록

The screenshot shows a web browser window for the Google Cloud Platform (GCP) free trial registration. The URL in the address bar is https://console.cloud.google.com/freetrial?_ga=1.16145334.1456799681.1498616668&pli=1&page=1. The page title is "Google Cloud Platform".

The main content area has a blue header bar with the text "Cloud Platform 무료로 사용해 보기" and the "Google" logo.

고객 정보

- 계정 유형**: 사업자 (Edit icon)
- 이름 및 주소**: 도/시 (Edit icon)
- 구/군** (Edit icon)
- 주소** (Edit icon)
- 주소 입력란 1** (Input field)
- 주소 이력란 2** (Input field)

Google

모든 Cloud Platform 제품에 액세스
앱, 웹사이트, 서비스를 구축하고 실행하는 데 필요한 Firebase, Google Maps API 등의 모든 기능을 이용할 수 있습니다.

무료 크레딧 300달러
가입하여 Google Cloud Platform에서 이후 12개월 동안 사용할 수 있는 300달러의 혜택을 받아 보세요.

무료 평가판 종료 후 자동 청구되지 않음
신용카드를 요청하는 이유는 자동 가입을 방지하기 위해서입니다. 유료 계정으로 직접 업그레이드하지 않는 한 요금이 청구되지 않습니다.

48 / 99

구글 클라우드 프로젝트 생성

The screenshot shows the Google Cloud Platform dashboard for the project 'konlper'. The left sidebar includes links for Home, Compute Engine, Datastore, Storage, SQL (selected), Spanner, StackDriver, Monitoring, Debugger, Metrics, Logs, and Reporting. The main content area displays the 'Compute Engine' section with a chart showing CPU usage over time, a table of resources (Compute Engine instance 1, Cloud SQL instance 1), and sections for Metrics, Cloud SQL, and Reporting.

Google Cloud Platform 대시보드

Compute Engine

프로젝트 정보

konlper

프로젝트 ID: konlper-168808
#636441884378

→ 프로젝트 설정으로 이동

리소스

Compute Engine
인스턴스 1개

Cloud SQL
인스턴스 1개

추적

지난 7일 동안에는 추적 데이터가 없습니다.

→ Stackdriver 추적 시작하기

Compute Engine

CPU(%)

6월 28일 오후 12:30 6월 28일 오후 1:06

CPU: 3.083

→ Compute Engine 대시보드로 이동

SQL

사용된 저장소(바이트)

1G
768M
512M

Google Cloud Platform 상태

모든 서비스 정상

→ Cloud 상태 대시보드로 이동

결제

\$0.00

현재까지 이번 달 예상 청구 금액

→ 청구 세부정보 보기

오류 보고

오류가 감지되지 않았습니다. 오류 보고를 설정하셨나요?

→ 오류 보고 설정 방법 알아보기

뉴스

R MySQL 연결

DBI로 MySQL을 연결하는 패키지와 사용

```
library(RMySQL)
con <- dbConnect(MySQL(),
                  user = user,
                  password = pw,
                  host = host,
                  dbname = "recom")
dbListTables(conn = con)
dbwriteTable(conn = con, name = 'tran', value = "./data/recomen/tran.csv")
dbReadTable(conn = con, name = "tran")
```

MySQL 인스턴스 생성

The screenshot shows the Google Cloud Platform interface for managing Cloud SQL instances. The top navigation bar includes the project name 'SQL 개요 - konlper' and the URL 'https://console.cloud.google.com/sql/instances?project=konlper-168808'. The main menu bar has 'Google Cloud Platform' and 'konlper' dropdowns, along with search and user icons. On the left, there's a sidebar with 'SQL' and '인스턴스' buttons. The main content area is titled 'Cloud SQL' and 'Cloud SQL 인스턴스'. It contains a descriptive paragraph about Cloud SQL instances and two buttons at the bottom: '인스턴스 생성' (Instance Creation) and '자세히 알아보기' (Learn More). The entire interface is in Korean.

엔진선택

The screenshot shows a web browser window for the Google Cloud Platform SQL service. The URL in the address bar is <https://console.cloud.google.com/sql/choose-instance-engine?project=konlper-168808>. The page title is "SQL - konlper". The main content area is titled "데이터베이스 엔진 선택" (Database Engine Selection). It offers two options: "MySQL" (selected) and "PostgreSQL" (beta). Both options mention specific versions: MySQL 5.6 또는 5.7 and PostgreSQL 9.6. A blue "다음" (Next) button is at the bottom.

데이터베이스 엔진 선택

MySQL
버전 5.6 또는 5.7

PostgreSQL 베타
버전 9.6

다음

MySQL 2세대 선택

The screenshot shows the Google Cloud Platform SQL pre-create interface. The URL in the browser is <https://console.cloud.google.com/sql/pre-create?project=konlper-168808>. The page title is "MySQL 인스턴스 유형 선택".

MySQL 2세대(권장)

저렴한 비용으로 높은 성능과 큰 저장용량이 제공됩니다.

- 1세대 처리량의 최대 7배 및 저장용량 크기 20배
- 대부분의 사용 사례에서 1세대보다 비용 낮음
- 고가용성 장애 조치를 추가하고 복제본을 읽을 수 있는 옵션 제공
- 백업 기간 및 유지관리 기간 구성 가능
- MySQL 5.6 및 5.7만 지원됩니다.

2세대 선택

MySQL 1세대(이전)

기본 성능과 저장용량 크기를 제공하는 이전 버전의 Cloud SQL입니다.
MySQL 5.7을 지원하지 않습니다.

1세대 선택

인스턴스 설정

The screenshot shows the Google Cloud Platform SQL creation interface. The top navigation bar includes the project name 'konlper' and the URL 'https://console.cloud.google.com/sql/create-instance-mysql?project=konlper-168808'. The main title is 'MySQL 2세대 인스턴스 만들기'.

인스턴스 ID
이후에 변경할 수 없습니다. 소문자, 숫자, 하이픈을 사용할 수 있으며 문자로 시작해야 합니다.
[Empty input field]

데이터베이스 버전
MySQL 5.7

위치
성능을 향상하려면 필요한 서비스와 가까운 위치에 데이터를 보관하세요.
지역: us-central1
영역: 자동 선택

머신 유형
성능을 향상하려면 가장 큰 표를 보관할 수 있는 메모리가 제공되는 머신 유형을 선택하세요.
db-n1-standard-1
vCPU: 1
메모리: 3.75GB
[Change button]

네트워크 처리량(MB/초)
250/2,000
[Progress bar]

저장소 유형
선택한 후에는 변경할 수 없습니다.

접속 허용 IP 설정

The screenshot shows the Google Cloud Platform SQL creation interface for creating a MySQL 2nd-generation instance. The page is titled "MySQL 2세대 인스턴스 만들기".

유지관리 기간 (Maintenance Window):
시스템 자동 선택 기간 (System automatically selects duration): dropdown menu.

유지관리 시점 (Maintenance Time):
자동 선택 (Automatic selection): dropdown menu.

루트 비밀번호 (Root Password):
루트 사용자의 비밀번호를 설정하세요. (Set the root user's password.)
생성 (Generate): button.
 비밀번호 없음 (No password): checkbox.

Cloud SQL 플래그 (Cloud SQL Flags):
+ 항목 추가 (Add item): button.

승인된 네트워크 (Approved Networks):
아래에 IPv4 주소를 추가하여 인스턴스에 연결할 네트워크를 승인하세요. 아래에 추가한 주소를 통해서만 네트워크가 승인됩니다. (Add an IPv4 address below to approve the network that can connect to the instance. Only networks with the added addresses will be approved.)
+ 네트워크 추가 (Add network): button.

접속 허용 IP 설정

The screenshot shows the Google Cloud Platform SQL interface for creating a MySQL 2nd Gen instance. The URL in the browser is <https://console.cloud.google.com/sql/create-instance-mysql?project=konlper-168808>.

The main form includes the following fields:

- 비밀번호 없음
- Cloud SQL 플래그: + 항목 추가
- 승인된 네트워크: 아래에 IPv4 주소를 추가하여 인스턴스에 연결할 네트워크를 승인하세요. 아래에 추가한 주소를 통해서만 네트워크가 승인됩니다.

A modal window titled "새 네트워크" (New Network) is open, containing the following fields:

- 이름 (선택사항): 없음
- 네트워크: CIDR 표기 사용 예: 199.27.25.0/24
- 완료, 취소

At the bottom of the modal is a button: + 네트워크 추가 (Add Network).

현재 IP 확인하기

A screenshot of a Google search results page. The search bar at the top contains the query "myip". Below the search bar, there are tabs for "전체" (All), "동영상" (Videos), "뉴스" (News), "이미지" (Images), "지도" (Maps), and "더보기" (More). There are also links for "설정" (Settings) and "도구" (Tools). The search results show approximately 20,800,000 results in 0.31 seconds. The first result is a link to "What Is My IP Address - IP Address Tools and Info - WhatIsMyIP.com ®" with the URL <https://www.whatismyip.com/>. The snippet for this result includes the text: "We provide IP address tools that allow users to perform an Internet Speed Test, IP address lookup, proxy detection, IP Whois Lookup, and more." Below this, there are two sections: "My IP Information" and "IP Address Lookup". The second result is "What Is My IP Address? IP Address Tools and More" with the URL whatismyipaddress.com/. The snippet for this result includes the text: "IP address lookup, location, proxy detection, email tracing, IP hiding tips, blacklist check, speed test, and forums. Find, get, and show my IP address." The third result is "What is my IP address? - IP Location" with the URL <https://www.iplocation.net/find-ip-address>. The snippet for this result includes the text: "This webpage displays public IP address of your computer or router assigned by your ISP." The fourth result is "What's My IP Address? | Online Privacy and Security Tool - ExpressVPN" with the URL <https://www.expressvpn.com/what-is-my-ip>. The snippet for this result includes the text: "IP address lookup. What do others know about your location? See the true IP address of your VPN or proxy server. Learn to hide your IP address in 2 minutes." The fifth result is "What's My IP Address? Networking Tools & More" with the URL www.whatsmyip.org/.

SQL - konlper x G myip - Google 검색 x

← → C 🔒 안전함 | https://www.google.co.kr/search?q=myip&oq=myip&aqs=chrome..69i57.1526j0j4&sourceid=chrome&ie=UTF-8

Google myip

전체 동영상 뉴스 이미지 지도 더보기 설정 도구

검색결과 약 20,800,000개 (0.31초)

What Is My IP Address - IP Address Tools and Info - WhatIsMyIP.com ®
<https://www.whatismyip.com/> 이 페이지 번역하기

We provide IP address tools that allow users to perform an Internet Speed Test, IP address lookup, proxy detection, IP Whois Lookup, and more.

My IP Information
My IP Information tool shows your ip address, city, state, country ...
[whatismyip.com 검색결과 더보기 »](#)

IP Address Lookup
The IP Address Lookup tool includes the following IP ...

What Is My IP Address? IP Address Tools and More
whatismyipaddress.com/ 이 페이지 번역하기

IP address lookup, location, proxy detection, email tracing, IP hiding tips, blacklist check, speed test, and forums. Find, get, and show my IP address.

What is my IP address? - IP Location
<https://www.iplocation.net/find-ip-address> 이 페이지 번역하기

This webpage displays public IP address of your computer or router assigned by your ISP.

What's My IP Address? | Online Privacy and Security Tool - ExpressVPN
<https://www.expressvpn.com/what-is-my-ip> 이 페이지 번역하기

IP address lookup. What do others know about your location? See the true IP address of your VPN or proxy server. Learn to hide your IP address in 2 minutes.

What's My IP Address? Networking Tools & More
www.whatsmyip.org/ 이 페이지 번역하기

Your IP Address plus Port Scanners, Traceroute, HTTP Compression Test, Ping, Whois, DNS, IP Geo

whatismyip

<https://www.whatismyip.com/>

The screenshot shows a web browser window with the URL <https://www.whatismyip.com/> in the address bar. The page displays the following information:

- WhatIsMyIP.com®** logo and tagline "THE IP ADDRESS EXPERTS".
- Navigation menu: IP Tools, How To, Resources, Home, Speed Test, IP Lookup, Hide My IP, Change My IP, Log In, Create Account.
- Advertisement banner: "Ads by Google" featuring "IP Location", "Check Your IP", and "My IP Address".
- IP Address: **112.217.214.251**.
- Geolocation details:
 - City: Seoul
 - State: Seoul-teukbyeolsi
 - Country: Korea, Republic of
 - ISP: LG Dacom Corporation
- Grammarly Instant Grammar Checker advertisement.
- Footer navigation: My IP Information, What My IP Says About Me, Proxy Check.
- Social media links: Facebook (42.9K), Twitter (1.4K).

ip 입력

The screenshot shows the Google Cloud Platform SQL interface for creating a MySQL 2nd-generation instance. The top navigation bar includes tabs for 'SQL - konlper' and 'What Is My IP Address'. The URL in the address bar is <https://console.cloud.google.com/sql/create-instance-mysql?project=konlper-168808>. The main content area is titled 'MySQL 2세대 인스턴스 만들기' (Create MySQL 2nd-gen instance). It features a checkbox for '비밀번호 없음' (No password) and a section for 'Cloud SQL 플래그' (Cloud SQL flags) with a '+ 항목 추가' (Add item) button. Below these, there's a section for '승인된 네트워크' (Approved networks) with a note about adding IPv4 addresses. A modal window titled '새 네트워크' (New network) is open, showing fields for '이름 (선택사항)' (Name (optional)) containing '없음' (None), '네트워크 CIDR 표기 사용' (Use CIDR notation for network), and the IP address '112.217.214.251'. Buttons for '완료' (Done) and '취소' (Cancel) are at the bottom of the modal, along with a '+ 네트워크 추가' (Add network) button.

root 계정 비밀번호 생성

The screenshot shows the Google Cloud Platform SQL creation interface for creating a MySQL 2nd-generation instance. The current step is generating the root password.

Root 비밀번호
루트 사용자의 비밀번호를 설정하세요. 자세히 알아보기
[비밀번호 편집] [생성]

비밀번호 없음

Cloud SQL 플래그
[항목 추가]

승인된 네트워크
아래에 IPv4 주소를 추가하여 인스턴스에 연결할 네트워크를 승인하세요. 아래에 주가 한 주소를 통해서만 네트워크가 승인됩니다.
112.217.214.251 [저장/제거]

[네트워크 추가]

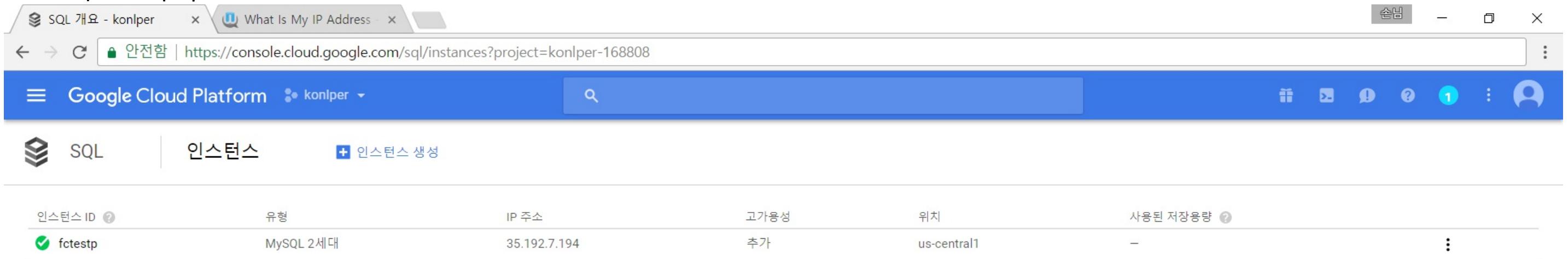
SQL 인스턴스 생성중

The screenshot shows the Google Cloud Platform interface for managing SQL instances. The top navigation bar includes tabs for 'SQL 개요 - konlper' and 'What Is My IP Address'. The URL in the address bar is <https://console.cloud.google.com/sql/instances?project=konlper-168808>. The main header features the 'Google Cloud Platform' logo, a search bar, and various navigation icons. Below the header, there are three main tabs: 'SQL' (selected), '인스턴스' (Instances), and '+ 인스턴스 생성' (Create Instance). The main content area displays a table of existing instances:

| 인스턴스 ID | 유형 | IP 주소 | 고가용성 | 위치 | 사용된 저장용량 |
|---------|-----------|-------|------|-------------|----------|
| fctestp | MySQL 2세대 | - | - | us-central1 | - |

RMySQL로 연결하기

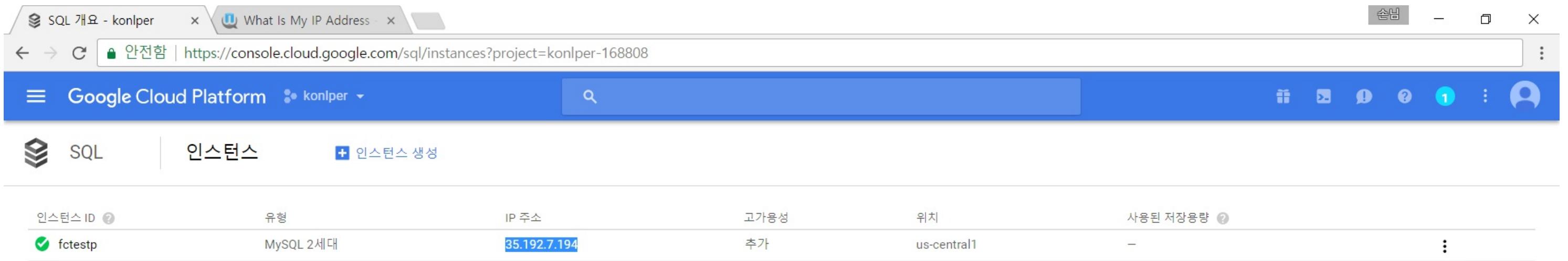
- host 주소 파악



The screenshot shows the Google Cloud Platform interface for managing SQL instances. The top navigation bar includes tabs for 'SQL 개요' and 'What Is My IP Address'. The main content area displays a single instance named 'fctestp' under the 'Instances' tab. The instance details are as follows:

| 인스턴스 ID | 유형 | IP 주소 | 고가용성 | 위치 | 사용된 저장용량 |
|---------|-----------|--------------|------|-------------|----------|
| fctestp | MySQL 2세대 | 35.192.7.194 | 추가 | us-central1 | - |

RMySQL로 연결하기



The screenshot shows the Google Cloud Platform interface for managing SQL instances. The top navigation bar includes tabs for 'SQL 개요' and 'What Is My IP Address'. The main title bar displays the URL <https://console.cloud.google.com/sql/instances?project=konlper-168808>. The header also shows the project name 'konlper' and a user profile icon.

The main content area is titled 'Google Cloud Platform' and 'konlper'. It features a search bar and several status indicators (green, yellow, red) for different services. Below this is a navigation menu with 'SQL' selected, followed by '인스턴스' and a button to '인스턴스 생성'.

A table lists the existing MySQL instance:

| 인스턴스 ID | 유형 | IP 주소 | 고가용성 | 위치 | 사용된 저장용량 |
|---------|-----------|--------------|------|-------------|----------|
| fctestp | MySQL 2세대 | 35.192.7.194 | 추가 | us-central1 | - |

데이터베이스 만들기

The screenshot shows the RStudio interface with the following components:

- File Explorer:** Shows the project structure under "dabrp_classnote2".
- Code Editor:** Displays an R script named "codeForClass2.R" containing the following code:

```
1 if (!require(devtools)) install.packages("devtools")
2 if (!require(DBI)) devtools::install_github("rstats-db/DBI")
3 if (!require(RSQLite)) devtools::install_github("rstats-db/RSQLite")
4 if (!require(RMySQL)) devtools::install_github("rstats-db/RMySQL")
5
6
7 # user<- "root"
8 # pw<- "XXXXXXXXXXXXXX"
9 # host<- 'XXX.XXX.XXX.XXX'
10
11 # save(user,pw,host,file = "./gsql.RData")
12
13 load("./gsql.RData")
14
15 library(RMySQL)
16 con <- dbConnect(MySQL(),
17                   user = user,
18                   password = pw,
19                   host = host,
20                   dbname = "test")
21 dbListTables(conn = con)
22 dbWriteTable(conn = con, name = 'Test', value = as.data.frame(iris))
23 dbReadTable(conn = con, name = "Test")
```

- Console:** Displays the R startup message and the command "R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"".
- Environment:** Shows the current session environment.
- Plots:** No plots are currently displayed.
- Packages:** No packages are currently displayed.
- Help:** No help pages are currently displayed.
- Viewer:** No files are currently displayed.

데이터베이스 만들기

The screenshot shows the Google Cloud Platform SQL Instances page. The instance 'fctestp' is selected. The 'Storage usage' section displays a graph of storage usage over time, with a legend indicating 'Storage usage: 1.14G'. The graph shows data points at 12:30, 12:45, 1:00, and 1:16 on June 28. Below the graph, there is a 'Logs' section with a 'View logs' button.

Storage usage graph details:

- Legend: Storage usage: 1.14G
- Time points: 6월 28일 오후 12:30, 6월 28일 오후 12:45, 6월 28일 오후 1:00, 6월 28일 오후 1:16

Logs section:

- Button: 오류 로그 보기 (View logs)

Bottom navigation:

- Cloud Shell을 사용해 연결 (Connect via Cloud Shell)
- 모든 연결 방법 보기 (View all connection methods)

데이터베이스 만들기

The screenshot shows the Google Cloud Platform SQL Instances page. The URL in the browser is <https://console.cloud.google.com/sql/instances/fctestp/databases?project=konlper-168808>. The page title is "fctestp 데이터베이스". The main content area is titled "인스턴스 세부정보" and shows the instance "fctestp" selected. Below it, the MySQL 2세대 database "fctestp" is listed under "MySQL 데이터베이스". A button labeled "데이터베이스 만들기" is highlighted with a red box. At the bottom, there is a table listing existing databases:

| 이름 | 문자 세트 | 대조 | 유형 |
|--------------------|-------|-----------------|-----|
| information_schema | utf8 | utf8_general_ci | 시스템 |
| mysql | utf8 | utf8_general_ci | 시스템 |
| performance_schema | utf8 | utf8_general_ci | 시스템 |

데이터베이스 만들기

The screenshot shows the Google Cloud Platform SQL interface for the 'fctestp' instance. The main page displays the MySQL 2세대 database 'fctestp' (Master, us-central1). A modal window titled '데이터베이스 만들기' (Create Database) is open, prompting for the database name. The input field contains 'MySQL 스키마 개체 이름 규칙을 준수해야 합니다.' (Must follow MySQL schema object name rules). Below the input field are dropdowns for character set ('utf8') and collation ('기본 대조'). At the bottom of the modal are '취소' (Cancel) and '만들기' (Create) buttons.

데이터베이스 만들기

The screenshot shows the Google Cloud Platform SQL interface for the 'fctestp' instance. The main view displays the MySQL 2세대 database 'fctestp' with three tables: 'information_schema', 'mysql', and 'performance_schema'. A modal window titled '데이터베이스 만들기' (Create Database) is open, prompting for a database name ('test'), character set ('utf8'), and collation ('기본 대조'). The '만들기' (Create) button is visible at the bottom right of the modal.

SQL

fctestp
마스터
us-central1

인스턴스 세부정보

수정 가져오기 내보내기 다시 시작 종지 삭제

MySQL 2세대

개요 액세스 제어 데이터베이스 백업 복제본 작업

MySQL 데이터베이스

데이터베이스 만들기

| 이름 | 문자 세트 | 대조 |
|--------------------|-------|-----------------|
| information_schema | utf8 | utf8_general_ci |
| mysql | utf8 | utf8_general_ci |
| performance_schema | utf8 | utf8_general_ci |

데이터베이스 만들기

데이터베이스 이름
MySQL 스키마 개체 이름 규칙을 준수해야 합니다.
test

문자 세트 대조 (선택사항)
utf8 기본 대조

취소 만들기

데이터베이스 만들기

The screenshot shows the Google Cloud Platform SQL interface for creating a new database. The main page displays the 'fctestp' instance details, including its status as a MySQL 2nd Generation master in the us-central1 region. The 'Databases' tab is selected, showing existing databases like 'information_schema', 'mysql', and 'performance_schema'. A modal dialog box titled '데이터베이스 만들기' (Create Database) is open, prompting for the database name ('test'), character set ('utf8'), and collation ('utf8'). The dialog also includes buttons for '취소' (Cancel), '만들기' (Create), and an ellipsis (...).

fctestp 데이터베이스 - k x What Is My IP Address - x

← → ⌛ 안전함 | https://console.cloud.google.com/sql/instances/fctestp/databases?project=konlper-168808

Google Cloud Platform konlper :

SQL

fctestp 마스터 us-central1

인스턴스 세부정보 ← 수정 가져오기 내보내기 다시 시작 종지 삭제

MySQL 데이터베이스

데이터베이스 만들기

| 이름 | 문자 세트 | 대조 |
|--------------------|-------|----|
| information_schema | utf8 | |
| mysql | utf8 | |
| performance_schema | utf8 | |

데이터베이스 이름
MySQL 스키마 개체 이름 규칙을 준수해야 합니다.
test

문자 세트 대조 (선택사항)
utf8 기본 대조

취소 만들기 ...

host와 dbname 설정

The screenshot shows the Google Cloud Platform SQL Instances page. The instance 'fctestp' is selected. The '데이터베이스' tab is active, displaying the MySQL databases:

| 이름 | 문자 세트 | 대조 | 유형 |
|--------------------|-------|-----------------|-----|
| information_schema | utf8 | utf8_general_ci | 시스템 |
| mysql | utf8 | utf8_general_ci | 시스템 |
| performance_schema | utf8 | utf8_general_ci | 시스템 |
| test | utf8 | utf8_general_ci | 사용자 |

A success message at the bottom left says "데이터베이스를 생성했습니다." (Database created successfully).

실습 진행

The screenshot shows the RStudio interface with the following components:

- File Explorer:** Shows the project structure under "dabrp_classnote2".
- Code Editor:** Displays an R script named "codeForClass2.R" containing code to install packages, connect to MySQL, and interact with the database.
- Console:** Shows the R startup message and standard disclaimer text.
- Environment:** Shows the current session environment and git status.

```
1 if (!require(devtools)) install.packages("devtools")
2 if (!require(DBI)) devtools::install_github("rstats-db/DBI")
3 if (!require(RSQLite)) devtools::install_github("rstats-db/RSQLite")
4 if (!require(RMySQL)) devtools::install_github("rstats-db/RMySQL")
5
6
7
8 # user<- "root"
9 # pw<- "XXXXXXXXXXXXXXXXXX"
10 # host<- 'XXX.XXX.XXX.XXX'
11
12 # save(user,pw,host,file = "./gsql.RData")
13
14 load("./gsql.RData")
15
16 library(RMySQL)
17 con <- dbConnect(MySQL(),
18                   user = user,
19                   password = pw,
20                   host = host,
21                   dbname = "test")
22 dbListTables(conn = con)
23 dbWriteTable(conn = con, name = 'Test', value = as.data.frame(iris))
24 dbReadTable(conn = con, name = "Test")
25
26
```

Console output:

```
R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

R bigquery 연결

`query_exec`가 첫 실행이면 브라우저에서 권한 확인을 합니다.

```
library(bigrquery)
project <- "konlper-168808"
sql <- "SELECT * FROM [konlper-168808:recom.channel] LIMIT 5"
query_exec(sql, project = project)
```

bigquery 인스턴스 생성

The screenshot shows the Google Cloud Platform (GCP) SQL Instances page. The URL in the browser is <https://console.cloud.google.com/sql/instances/fctestp/databases?project=konlper-168808>. The page title is "fctestp 데이터베이스". The main content area is titled "인스턴스 세부정보" and shows a MySQL instance named "fctestp". Below the instance name, it says "MySQL 2세대". There are tabs for "주요", "액세스 제어", "데이터베이스", "백업", "복제본", and "작업". The "데이터베이스" tab is selected. A sub-section titled "MySQL 데이터베이스" shows a list of databases:

| 이름 | 문자 세트 | 대조 | 유형 |
|--------------------|-------|-----------------|-----|
| information_schema | utf8 | utf8_general_ci | 시스템 |
| mysql | utf8 | utf8_general_ci | 시스템 |
| performance_schema | utf8 | utf8_general_ci | 시스템 |
| recom | utf8 | utf8_general_ci | 사용자 |
| test | utf8 | utf8_general_ci | 사용자 |

On the left sidebar, under "제품", the "Compute Engine" section is expanded, showing "fctestp" as a MySQL 2세대 instance. Other products listed include Container Registry, Docker Registry, BigQuery, Dataflow, DataProc, and ML Engine.

bigquery 첫 화면

The screenshot shows the Google BigQuery web interface. At the top, there's a navigation bar with tabs for 'fcitestp 데이터베이스' and 'Google BigQuery'. Below the navigation bar is a header with a search icon, a lock icon indicating a secure connection, and the URL 'https://bigquery.cloud.google.com/welcome/konlper-168808'. The main content area has a title 'Welcome to BigQuery!' followed by a brief introduction: 'Google BigQuery is a web service that lets you do interactive analysis of massive datasets—up to billions of rows. Scalable and easy to use, BigQuery lets developers and businesses tap into powerful data analytics on demand.' Below the introduction, there's a section titled 'To get started, try one of the following options:' with a bulleted list:

- Read our [BigQuery Quickstart guide](#)
- Run a query against our sample data by clicking "Compose Query"
- Create a new dataset and load some of your own data into a table using the menu on the left
- Learn about [cost control options in BigQuery](#)
- For more information on the UI, see the [BigQuery Web UI guide](#)

On the left side, there's a sidebar with a red 'COMPOSE QUERY' button. Below it are 'Query History' and 'Job History' links. A 'Filter by ID or label' input field contains 'konlper' with a dropdown arrow. Below the input field, it says 'No datasets found in this project.' and 'Please create a dataset or select a new project from the menu above.' Under the 'Public Datasets' heading, there's a list of datasets: 'bigquery-public-data:hacker_news', 'bigquery-public-data:noaa_gsod', 'bigquery-public-data:samples', 'bigquery-public-data:usa_names', 'gdelt-bq:hathitrustbooks', 'gdelt-bq:internetarchivebooks', 'lookerdata:cdc', 'nyc-tlc:green', and 'nyc-tlc:yellow'.

dataset 만들기

The screenshot shows the Google BigQuery web interface. At the top, there's a navigation bar with tabs for 'fcitestp 데이터베이스' and 'Google BigQuery'. The URL in the address bar is <https://bigquery.cloud.google.com/welcome/konlper-168808>. On the right side of the header are icons for help, notifications, and a user profile.

The main content area has a title 'Welcome to BigQuery!' and a brief introduction: 'Google BigQuery is a web service that lets you do interactive analysis of massive datasets—up to billions of rows. Scalable and easy to use, BigQuery lets developers and businesses tap into powerful data analytics on demand.' Below this, there's a section titled 'To get started, try one of the following options:' with two bullet points: 'Read our [BigQuery Quickstart guide](#)' and 'Run a query against our sample data by clicking "Compose Query"'. A dropdown menu is open over this section, containing options like 'Create new dataset', 'Switch to project', and 'Refresh'. To the left of the main content, there's a sidebar with a red 'COMPOSE QUERY' button, 'Query History', 'Job History', a filter input 'Filter by ID or label', and a dropdown menu showing 'konlper' and its contents: 'No datasets found in this project.', 'Please create a dataset or select a new project from the menu above.', and a 'Public Datasets' section listing various datasets like 'bigquery-public-data:hacker_news', 'bigquery-public-data:noaa_gsod', etc.

Public Datasets

- ▶ [bigquery-public-data:hacker_news](#)
- ▶ [bigquery-public-data:noaa_gsod](#)
- ▶ [bigquery-public-data:samples](#)
- ▶ [bigquery-public-data:usa_names](#)
- ▶ [gdelt-bq:hathitrustbooks](#)
- ▶ [gdelt-bq:internetarchivebooks](#)
- ▶ [lookerdata:cdc](#)
- ▶ [nyc-tlc:green](#)
- ▶ [nyc-tlc:yellow](#)

fctestp 데이터베이스 - k x Google BigQuery x

← → C 🔒 안전함 | https://bigquery.cloud.google.com/welcome/konlper-168808

Google BigQuery

?

c

COMPOSE QUERY

Query History

Job History

Filter by ID or label



konlper



No datasets found in this project.

Please create a dataset or select a new project from the menu above.

▼ Public Datasets

- ▶ bigquery-public-data:hacker_news
- ▶ bigquery-public-data:noaa_gsod
- ▶ bigquery-public-data:samples
- ▶ bigquery-public-data:usa_names
- ▶ gdelt-bq:hathitrustbooks
- ▶ gdelt-bq:internetarchivebooks
- ▶ lookerdata:cdc
- ▶ nyc-tlc:green
- ▶ nyc-tlc:yellow

Welcome to BigQuery!

Google BigQuery is a web service that lets you do interactive analysis of massive datasets—up to billions of rows. Scalable and easy to use, BigQuery lets developers and businesses tap into powerful data analytics on demand.

To get started, try one of the following options:

- Read our [BigQuery Quickstart guide](#)
- Run a query against our sample data by clicking "Compose Query"
- Create a new dataset and load some of your own data into a table using the menu on the left
- Learn about [BigQuery](#)
- For more information, see the [BigQuery documentation](#)

Create Dataset

Dataset ID ?

Data location ?

Data expiration Never In days. ?

OK Cancel



table 생성

The screenshot shows the Google BigQuery web interface for creating a new table. The URL in the browser is <https://bigquery.cloud.google.com/createtable/konlper-168808:recom>.

Source Data: The "Create from source" option is selected. Under "Repeat job", "Select Previous Job" is chosen. The "Location" dropdown is set to "File upload" and "Choose file" is selected, showing "No file chosen". The "File format" dropdown is set to "CSV".

Destination Table: The "Table name" dropdown is set to "recom". The "Table type" dropdown is set to "Native table".

Schema: The "Name" field is empty. The "Type" dropdown is set to "STRING" and the "Mode" dropdown is set to "NULLABLE". A button "Edit as Text" is visible.

Options: The "Field delimiter" is set to "Comma". The "Header rows to skip" is set to "0". The "Number of errors allowed" is set to "0". The "Allow quoted newlines" and "Allow jagged rows" options are both unchecked.

큰 데이터는 storage를 통해 업로드

The screenshot shows the 'Create Table' page in the Google BigQuery web UI. The URL in the browser is <https://bigquery.cloud.google.com/createtable/konlper-168808:recom>. The left sidebar includes 'COMPOSE QUERY', 'Query History', 'Job History', a filter for 'konlper', and a list of 'Public Datasets'. The main form is titled 'Create Table' and has the following fields:

- Source Data:** 'Create from source' is selected.
- Repeat job:** 'Select Previous Job' is shown.
- Location:** 'File upload' is selected, and a file named 'tran.csv (1484069730 bytes)' is chosen.
- File format:** 'CSV' is selected.
- Destination Table:** 'Table name' is set to 'recom' and 'Destination table name' is also 'recom'. 'Table type' is 'Native table'.
- Schema:** 'Automatically detect' is checked.
- Options:** 'Field delimiter' is 'Comma', 'Header rows to skip' is '0', and 'Number of errors allowed' is '0'.

A note at the bottom of the location section states: "Uploads from the BigQuery web UI are limited to 10 MB. For larger data sizes, please load data from Google Cloud Storage."

버킷 만들기

The screenshot shows the Google Cloud Platform Storage interface for creating a new bucket. The browser title bar indicates the user is on the 'Bucket 생성 - konlper' tab within the 'Google BigQuery' project. The URL in the address bar is <https://console.cloud.google.com/storage/create-bucket?project=konlper-168808>. The main content area is titled '버킷 만들기' (Create Bucket) and shows the following configuration:

- Storage** tab selected.
- Bucket Name:** yarling-stratum-4912
- Basic Storage Class:** Multi-Regional (selected).
 - Description: 동영상을 스트리밍하고 인기 웹 콘텐츠를 호스팅하는 데 사용하세요. 전 세계에서 자주 액세스하는 데이터에 적합합니다.
 - Regional**: 데이터를 저장하고 데이터 분석을 실행하는 데 사용하세요. 특정 지역에서 자주 액세스하는 데이터에 적합합니다.
 - Nearline**: 자주 액세스하지 않는 문서를 보관하는 데 사용하세요. 액세스 빈도가 한 달에 한 번 미만인 데이터에 적합합니다.
 - Coldline**: 거의 액세스하지 않는 문서를 보관하는 데 사용하세요. 액세스 빈도가 1년에 한 번 미만인 데이터에 적합합니다.
- Multi-Regional Location:** 미국 (selected).
- Labels:** 라벨 지정 (Labeling) section is collapsed.
- Buttons:** '만들기' (Create) button is highlighted in blue, and '취소' (Cancel) button is white.

버킷 생성 - konlper x Google BigQuery x

안전함 | https://console.cloud.google.com/storage/create-bucket?project=konlper-168808

Google Cloud Platform konlper :

Storage ← 버킷 만들기

브라우저 이름 ?
Cloud Storage 전체에서 고유해야 합니다. 개인정보 보호: 버킷 이름에 민감한 정보를 포함해서는 안 됩니다. 다른 사용자가 사용하려고 시도한 버킷 이름과 일치할 경우 이 사용자에게 버킷 이름이 노출될 수 있습니다.
yarling-stratum-4912

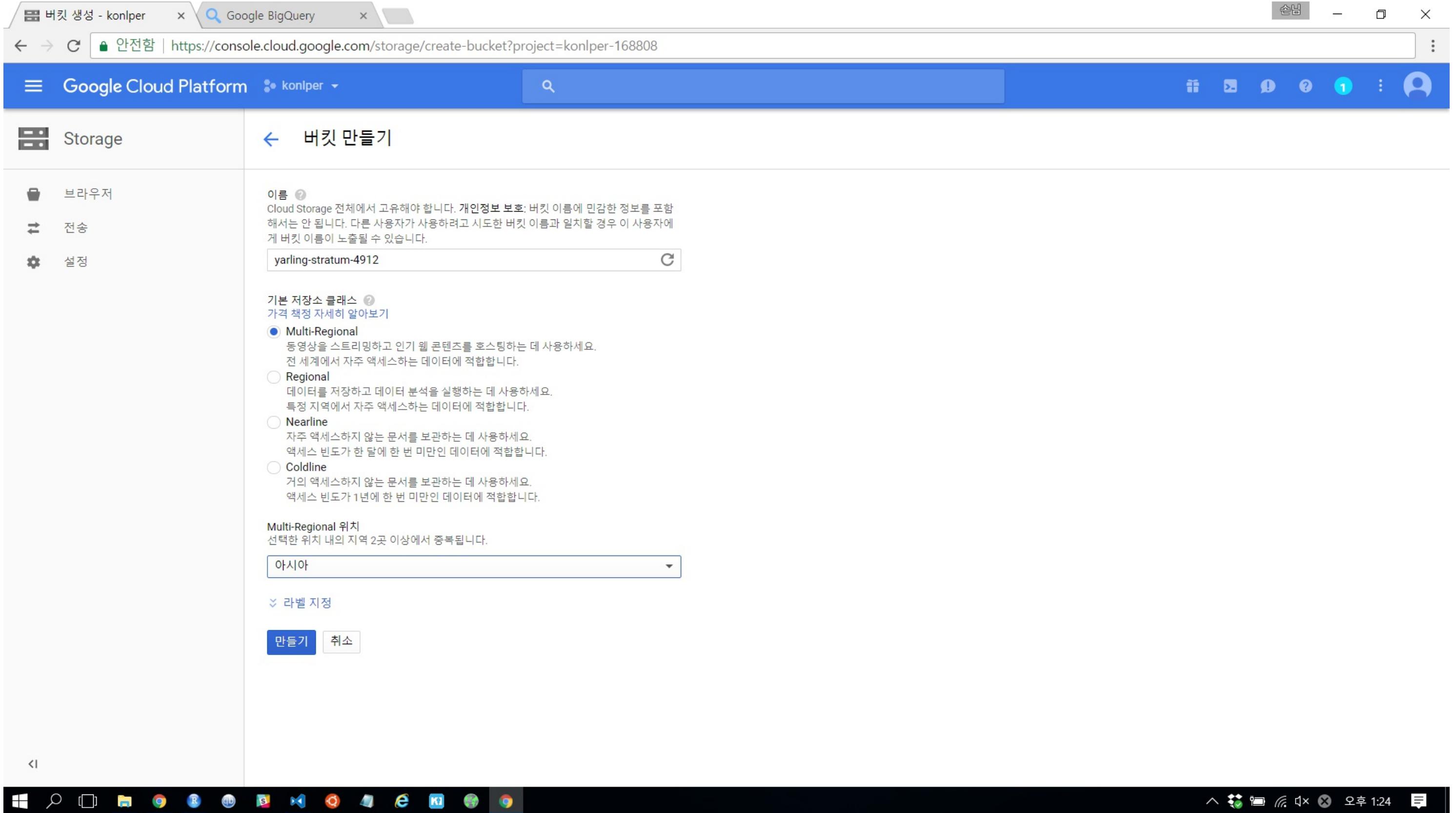
전송 설정

기본 저장소 클래스 ?
가격 책정 자세히 알아보기
 Multi-Regional
동영상을 스트리밍하고 인기 웹 콘텐츠를 호스팅하는 데 사용하세요.
전 세계에서 자주 액세스하는 데이터에 적합합니다.
 Regional
데이터를 저장하고 데이터 분석을 실행하는 데 사용하세요.
특정 지역에서 자주 액세스하는 데이터에 적합합니다.
 Nearline
자주 액세스하지 않는 문서를 보관하는 데 사용하세요.
액세스 빈도가 한 달에 한 번 미만인 데이터에 적합합니다.
 Coldline
거의 액세스하지 않는 문서를 보관하는 데 사용하세요.
액세스 빈도가 1년에 한 번 미만인 데이터에 적합합니다.

Multi-Regional 위치
선택한 위치 내의 지역 2곳 이상에서 중복됩니다.
아시아

▼ 라벨 지정

만들기 취소



버킷 확인

The screenshot shows the Google Cloud Platform Storage browser interface. The left sidebar is titled 'Storage' and has three options: '브라우저' (selected), '전송', and '설정'. The main area is titled '브라우저' and contains a search bar with the placeholder '프리픽스로 필터링...'. Below it is a table with one row, representing a bucket named 'yarling-stratum-4912'. The columns in the table are '이름' (Name), '기본 저장소 클래스' (Default storage class), '위치' (Location), and '라벨' (Labels). The bucket name is 'yarling-stratum-4912', the default storage class is 'Multi-Regional', the location is 'ASIA', and there are no labels.

| 이름 | 기본 저장소 클래스 | 위치 | 라벨 |
|----------------------|----------------|------|----|
| yarling-stratum-4912 | Multi-Regional | ASIA | |

폴더 업로드

The screenshot shows the Google Cloud Platform Storage browser interface. The top navigation bar includes tabs for 'yarling... 버킷 - konlper' and 'Google BigQuery'. The address bar shows a secure connection to 'https://console.cloud.google.com/storage/browser/yarling-stratum-4912/?project=konlper-168808'. The main header features the 'Google Cloud Platform' logo, a search bar, and various navigation icons.

The left sidebar is titled 'Storage' and contains three items: '브라우저' (selected), '전송', and '설정'. The '브라우저' item has a sub-menu with '프리픽스로 필터링...' and '버킷 / yarling-stratum-4912'. A message at the bottom of this menu states '이 버킷에 개체가 없습니다.'

The main content area is titled '브라우저' and includes buttons for '파일 업로드' (File Upload), '폴더 업로드' (Folder Upload), '폴더 생성' (Create Folder), '새로고침' (Refresh), '공개적으로 공유하기' (Share publicly), and '삭제' (Delete). There is also a search bar labeled '프리픽스로 필터링...'.

업로드 확인

The screenshot shows the Google Cloud Platform Storage browser interface. The left sidebar has 'Storage' selected, with '브라우저' (Browser) highlighted. The main area shows a single folder named 'recomen/'. The table below provides details about this folder.

| 이름 | 크기 | 유형 | 저장소 클래스 | 최종 수정 시간 | 공개적으로 공유하기 |
|----------|----|----|---------|----------|------------|
| recomen/ | - | 폴더 | - | - | |

업로드 확인

The screenshot shows the Google Cloud Platform Storage browser interface. The left sidebar has 'Storage' selected, with '브라우저' (Browser) highlighted. The main area shows a list of files in the bucket 'yarling-stratum-4912 / recomen'. The files are:

| 이름 | 크기 | 유형 | 저장소 클래스 | 최종 수정 시간 | 공개적으로 공유하기 |
|----------------|----------|--------------------------|----------------|--------------------|--------------------------|
| chennel.csv | 184.7KB | application/vnd.ms-excel | Multi-Regional | 17. 6. 28. 오후 1:24 | <input type="checkbox"/> |
| competitor.csv | 522.51KB | application/vnd.ms-excel | Multi-Regional | 17. 6. 28. 오후 1:24 | <input type="checkbox"/> |
| customer.csv | 454.13KB | application/vnd.ms-excel | Multi-Regional | 17. 6. 28. 오후 1:24 | <input type="checkbox"/> |
| item.csv | 192.9KB | application/vnd.ms-excel | Multi-Regional | 17. 6. 28. 오후 1:24 | <input type="checkbox"/> |
| membership.csv | 177.42KB | application/vnd.ms-excel | Multi-Regional | 17. 6. 28. 오후 1:24 | <input type="checkbox"/> |

At the top, there are buttons for '파일 업로드' (File Upload), '폴더 업로드' (Folder Upload), '폴더 생성' (Create Folder), '새로고침' (Refresh), '공개적으로 공유하기' (Share publicly), and '삭제' (Delete). The address bar shows the URL: https://console.cloud.google.com/storage/browser/yarling-stratum-4912/recomen/?project=konlper-168808.

gs:// 경로확인

The screenshot shows the Google BigQuery web interface with the title "Create Table".

Source Data: "Create from source" is selected. "Repeat job" is set to "Select Previous Job". "Location" is set to "Google Cloud Storage" with the path "gs://yarling-stratum-4912/recomen/chennel.csv". "File format" is set to "CSV".

Destination Table: "Table name" is "recom . chennel". "Table type" is "Native table".

Schema: "Automatically detect" is checked. A note says "Schema will be automatically generated."

Options: "Number of errors allowed" is "0". "Write preference" is "Write if empty".

Buttons: "Create Table" button at the bottom left.

작업 결과 확인

The screenshot shows the Google BigQuery web interface. The top navigation bar includes tabs for 'yarling... 버킷 - konlper' and 'yarling... 버킷'. The URL in the address bar is <https://bigquery.cloud.google.com/jobs/konlper-168808>. The main content area is titled 'Recent Jobs'.

COMPOSE QUERY button is visible in the top left.

Job History section is selected, showing a list of recent jobs:

- konlper** dataset:
 - recom** table
 - chennel** table
- Public Datasets**:
 - bigquery-public-data:hacker_news
 - bigquery-public-data:noaa_gsod
 - bigquery-public-data:samples
 - bigquery-public-data:usa_names
 - gdelt-bq:hathitrustbooks
 - gdelt-bq:internetarchivebooks
 - lookerdata:cdr
 - nyc-tlc:green
 - nyc-tlc:yellow

Recent Jobs table:

| Job Status | Job Description | Action Buttons | Last Updated |
|------------|---|-----------------|--------------|
| ✓ Load | gs://yarling-stratum-4912/recomen/chennel.csv to konlper-168808:recom.chennel | Repeat Load Job | 1:34PM |
| ✓ Load | gs://yarling-stratum-4912/recomen/chennel.csv to konlper-168808:recom.chennel | Repeat Load Job | 1:34PM |

query 실행화면

The screenshot shows the Google BigQuery web interface. On the left, there's a sidebar with 'COMPOSE QUERY' and sections for 'Query History' and 'Job History'. A dropdown menu shows 'konlper' selected, with 'recom' and 'chennel' as options. Below that is a 'Public Datasets' section listing various datasets like 'bigquery-public-data:hacker_news' and 'nyc-tlc:green'. The main area is titled 'New Query' and contains the following SQL code:

```
1 SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100
```

A green validation message at the bottom says 'Valid: This query will process 259 KB when run.' Below the query are buttons for 'RUN QUERY' (which has a checked checkbox), 'Save Query', 'Save View', 'Format Query', and 'Show Options'. The 'Results' tab is selected, and the status is 'Query Failed'. The error message is: 'Error: Encountered " "SELECT" "SELECT "" at line 1, column 1. Was expecting: <EOF>''. The job ID is listed as 'Job ID: konlper-168808:bquijob_3ada8983_15cecfbea1b'. A note at the bottom says 'Note: You can also find errors in your queries before running them. Click the ! below the query composition box to enable real-time validation.' At the very bottom, the same SQL code is shown again.

비용 발생 경고

The screenshot shows the Google BigQuery web interface. On the left, there's a sidebar with 'COMPOSE QUERY' and sections for 'Query History' and 'Job History'. Below that is a 'Filter by ID or label' dropdown set to 'konlper' which has 'recom' and 'chennel' under it. Under 'Public Datasets', there's a list of datasets like 'bigquery-public-data:hacker_news', 'bigquery-public-data:noaa_gsod', etc. The main area shows a 'New Query' window with the following SQL code:

```
1 SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100
```

Below the code, there's a 'Valid: This query will process 259' message, a 'RUN QUERY' button, and tabs for 'Results', 'Explanation', and 'Job Information'. A modal dialog titled 'Confirm query' is open, containing the following text:

With this query, you will be billed for all the data in the table (even if your query contains a LIMIT clause). If you're using the free tier, this query still counts against your free quota.

You can use table preview instead to see records for free and without affecting quotas.

Don't show this again

At the bottom of the modal are buttons for 'Run query', 'Go to table preview', and 'Cancel'. At the very bottom of the page, there's a note: 'Note: You can also find errors in your queries before running them. Click the ! below the query composition box to enable real-time validation.'

query 결과 확인

The screenshot shows the Google BigQuery interface. On the left, there's a sidebar with 'COMPOSE QUERY' and sections for 'Query History' and 'Job History'. Below that is a dropdown for 'konlper' with options 'recom' and 'chennel'. Under 'Public Datasets', a list includes 'bigquery-public-data:hacker_news', 'bigquery-public-data:noaa_gsod', 'bigquery-public-data:samples', 'bigquery-public-data:usa_names', 'gdelt-bq:hathitrustbooks', 'gdelt-bq:internetarchivebooks', 'lookerdata:cdc', 'nyc-tlc:green', and 'nyc-tlc:yellow'. The main area shows a 'New Query' window with the SQL command:

```
1 | SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100
```

A green validation message says 'Valid: This query will process 259 KB when run.' Below the query are buttons for 'RUN QUERY', 'Save Query', 'Save View', 'Format Query', and 'Show Options'. A status message indicates 'Query complete (2.1s elapsed, 259 KB processed)' with a checkmark icon. At the bottom, there are tabs for 'Results', 'Explanation', and 'Job Information', along with download options: 'Download as CSV', 'Download as JSON', 'Save as Table', and 'Save to Google Sheets'. The results table has columns: Row, cusID, chennel, and useCnt. The data shows 9 rows where all entries have a 'useCnt' of 1. The last row is highlighted.

| Row | cusID | chennel | useCnt |
|-----|-------|--------------|--------|
| 1 | 14 | A_MOBILE/APP | 1 |
| 2 | 74 | A_MOBILE/APP | 1 |
| 3 | 241 | A_MOBILE/APP | 1 |
| 4 | 304 | A_MOBILE/APP | 1 |
| 5 | 326 | A_MOBILE/APP | 1 |
| 6 | 377 | A_MOBILE/APP | 1 |
| 7 | 448 | A_MOBILE/APP | 1 |
| 8 | 518 | A_MOBILE/APP | 1 |
| 9 | 549 | A_MOBILE/APP | 1 |

Page navigation at the bottom includes 'Table' and 'JSON' buttons, and links for 'First', '< Prev', 'Rows 1 - 9 of 100', 'Next >', and 'Last'.

공개 데이터셋

The screenshot shows a browser window with four tabs open:

- yarling... 버킷 - konlper
- Google BigQuery
- yarling... 버킷
- Google Cloud Platform | Google BigQuery Public

The main content area is the "Documentation" page for Google BigQuery Public Datasets. The URL in the address bar is <https://cloud.google.com/bigquery/public-data/>.

The page includes a sidebar with links to various datasets:

- All Resources
 - Pricing and Quotas
 - Release Notes
 - Support
 - Public Datasets
 - 개요
 - 1000 Cannabis Genomes Project
 - Bay Area Bike Share Trips Data
 - Chicago Crime Data
 - Chicago Taxi Trips
 - EPA Historical Air Quality Data
 - GDELT Books Corpus
 - GitHub Data
 - Hacker News
 - Healthcare Common Procedure Coding System (HCPCS) Level II
 - IRS 990 Data
 - Major League Baseball
 - Medicare
 - NHTSA Traffic Fatality Data
 - NOAA GHCN Weather
 - NOAA GSOD Weather
 - NOAA ICOADS
 - NYC 311 Service Requests
 - NYC Citi Bike Trips

wikipedia dataset

The screenshot shows the Google BigQuery web interface. On the left, there's a sidebar with navigation links like 'COMPOSE QUERY', 'Query History', 'Job History', and a 'Filter by ID or label' dropdown set to 'konlper'. Below it is a 'Public Datasets' section listing various datasets such as 'bigquery-public-data:hacker_news', 'bigquery-public-data:noaa_gsod', and 'bigquery-public-data:samples'. The 'samples' dataset is currently selected, indicated by a red vertical bar.

In the main area, a 'New Query' window is open with the following SQL query:

```
1 SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100
```

To the right of the query, there are tabs for 'Query Editor' and 'UDF Editor', and a 'SQL' button. Below the editor are buttons for 'Query Table', 'Copy Table', 'Export Table', and 'Delete Table'.

A 'Create Dataset' modal is prominently displayed in the center. It contains fields for 'Dataset ID' (set to 'test'), 'Data location' (set to '(unspecified)'), and 'Data expiration' (set to 'In 5 days'). There are 'OK' and 'Cancel' buttons at the bottom of the modal.

At the bottom right of the interface, there's a page number '92 / 99'.

dataset 복사

The screenshot shows the Google BigQuery web interface. On the left, there's a sidebar with navigation links like 'COMPOSE QUERY', 'Query History', 'Job History', and a 'konlper' project dropdown. Below that is a 'Public Datasets' section with links to various datasets. A red vertical bar highlights the 'wikipedia' dataset under 'Public Datasets'. A context menu for this dataset is open, showing options: 'Copy table' (which is selected), 'Export table', and 'Delete table'. The main area shows a 'New Query' window with a single line of SQL: 'SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100'. Below the query is a green validation message: 'Valid: This query will process 259 KB when run.' At the bottom of the query window, there are buttons for 'RUN QUERY', 'Save Query', 'Save View', 'Format Query', and 'Show Options'. A status message says 'Query complete (2.1s elapsed, 259 KB processed)' with a checkmark icon. To the right of the query window is a 'Table Details: wikipedia' section. It shows three columns: 'Schema', 'Details', and 'Preview'. Under 'Schema', there are three rows for 'title' (STRING, REQUIRED), 'id' (INTEGER, NULLABLE), and 'language' (STRING, REQUIRED). The 'language' row has a note: 'Empty in the current dataset.' Under 'Details', there are two rows: 'title' with a note about namespaces, and 'id' with a note about article revision order. Under 'Preview', there is a note: 'Wikipedia segments its pages into namespaces (e.g. "Talk", "User", etc.)' followed by a list of constants: MEDIA = 202; // -2 in WP XML, but these values must be >0, SPECIAL = 201; // -1 in WP XML, but these values must be >0, MAIN = 0;, TALK = 1;, USER = 2;, USER_TALK = 3;. At the bottom right of the interface, there's a page number '93 / 99'.

Google BigQuery

COMPOSE QUERY

New Query

Query History

Job History

Filter by ID or label

konlper

recom

chennel

test

Public Datasets

bigrquery-public-data:hacker_news

bigrquery-public-data:noaa_gsod

bigrquery-public-data:samples

github_nested

github_timeline

gsod

natality

shakespeare

trigrams

wikipedia

bigrquery-public-data:usa_names

gdelt-bq:hathitrustbooks

gdelt-bq:internetarchivebooks

lockheeddata:ads

Compose Query

SQL

1 | SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100

Valid: This query will process 259 KB when run.

RUN QUERY Save Query Save View Format Query Show Options

Query complete (2.1s elapsed, 259 KB processed)

Table Details: wikipedia

Schema Details Preview

title STRING REQUIRED The title of the page, as displayed on the page (not in the URL). Always starts with a capital letter and may begin with a namespace (e.g. "Talk:", "User:", "User Talk:", ...)

id INTEGER NULLABLE A unique ID for the article that was revised. These correspond to the order in which articles were created, except for the first several thousand IDs, which are issued in alphabetical order.

language STRING REQUIRED Empty in the current dataset.

Copy table Export table Delete table

Wikipedia segments its pages into namespaces (e.g. "Talk", "User", etc.)

MEDIA = 202; // -2 in WP XML, but these values must be >0
SPECIAL = 201; // -1 in WP XML, but these values must be >0
MAIN = 0;
TALK = 1;
USER = 2;
USER_TALK = 3;

93 / 99

yarling... 버킷 - konlper x Google BigQuery x yarling... 버킷 x Google Cloud Platform x Google BigQuery Public x NOAA Global Surface S... x

안전함 | https://bigquery.cloud.google.com/table/bigquery-public-data:samples.wikipedia

Google BigQuery

COMPOSE QUERY

New Query ?

Query History

Job History

Filter by ID or label ?

konlper

- recom
- chennel
- test

Public Datasets

- bigquery-public-data:hacker_news
- bigquery-public-data:noaa_gsod
- bigquery-public-data:samples
- github_nested
- github_timeline
- gsod
- nativity
- shakespeare
- trigrams
- wikipedia
- bigquery-public-data:usa_names
- gdelt-bq:hathitrustbooks
- gdelt-bq:internetarchivebooks
- lookerdata:cdr
- nyc-tlc:green

1 SELECT * FROM [konlper-168808:recom.chennel] LIMIT 100

Query Editor UDF Editor X SQL

Copy Table

Valid: This query is valid.

RUN QUERY

Destination project: konlper (konlper-168808)

Destination dataset: test

Destination table: wikipedia_copy

OK Cancel

Table Data

Schema

| title | id | language | wp_namespace | is_redirect |
|-----------------------------|--|-------------------------------|---|--|
| User..., User talk..., ...) | A unique ID for the article that was revised. These correspond to the order in which articles were created, except for the first several thousand IDs, which are issued in alphabetical order. | Empty in the current dataset. | Wikipedia segments its pages into namespaces (e.g. "Talk", "User", etc.) MEDIA = 202; // =-2 in WP XML, but these values must be >0 SPECIAL = 201; // =-1 in WP XML, but these values must be >0 MAIN = 0; TALK = 1; USER = 2; USER TALK = 2; | Versions later than ca. 200908 may have a redirection marker in the XML. |

Query Table Copy Table Export Table Delete Table



데이터 구조 보기

The screenshot shows the Google BigQuery web interface. On the left, there's a sidebar with navigation links like 'COMPOSE QUERY', 'Query History', 'Job History', and sections for 'konlper' (containing 'recom', 'chennel', and 'test') and 'Public Datasets' (listing various datasets from Google Cloud Platform). The main area is titled 'New Query' and contains a SQL code editor with the following query:

```
1 | SELECT * FROM [konlper-168808:test.wikipedia_copy] LIMIT 100
```

Below the query, a green bar indicates: 'Valid: This query will process 35.7 GB when run.' At the bottom of the query editor are buttons for 'RUN QUERY', 'Save Query', 'Save View', 'Format Query', and 'Show Options'. A checkmark icon is visible next to the 'Show Options' button.

On the right side, there's a detailed view of the schema for the 'wikipedia_copy' table:

| Field | Type | Nullable | Description |
|----------------------|---------|----------|--|
| contributor_id | INTEGER | NULLABLE | Typically, either (_id and _username) or _ip will be set. A (very) small fraction of edits have neither _ip or (_id and _username). They show up on Wikipedia as "(Username or IP removed)". |
| contributor_username | STRING | NULLABLE | Typically, either (_id and _username) or _ip will be set. A (very) small fraction of edits have neither _ip or (_id and _username). They show up on Wikipedia as "(Username or IP removed)". |
| timestamp | INTEGER | REQUIRED | In Unix time, seconds since epoch. |
| is_minor | BOOLEAN | NULLABLE | Corresponds to the "Minor Edit" checkbox on Wikipedia's edit page. |
| is_bot | BOOLEAN | NULLABLE | A special flag that some of Wikipedia's more active bots voluntarily set. |
| reversion_id | INTEGER | NULLABLE | If this edit is a reversion to a previous edit, this field records the revision_id that was reverted to. If the same article text occurred multiple times, then this will point to the earliest revision. Only revisions with greater than fifty characters are considered for this field. This is to avoid labeling multiple blankings as reversions. |
| comment | STRING | NULLABLE | Optional user-supplied description of the edit. Section edits are, by default, prefixed with /* Section Name */. |
| num_characters | INTEGER | REQUIRED | The length of the article after the revision was applied. |

At the bottom right, there's a page number '95 / 99'.

큰 데이터 query 진행

```
select title,sum(num_characters) as num_characters
from [konlper-168808:test.wikipedia_copy]
where regexp_match(title,'[ss]eoul')
group by title
order by num_characters desc;
```

query 결과

The screenshot shows the Google BigQuery interface with a query results page. The query is:

```
1 select title,sum(num_characters) as num_characters
2 from [konlper-168808:test.wikipedia_copy]
3 where regexp_match(title,'[Ss]eoul')
4 group by title
5 order by num_characters desc;
```

The results table shows the top 10 Korean titles and their character counts:

| Row | title | num_characters |
|-----|----------------------------|----------------|
| 1 | Seoul | 104656429 |
| 2 | User talk:JohnnySeoul | 78435528 |
| 3 | FC Seoul | 44944412 |
| 4 | Seoul National University | 18422979 |
| 5 | Talk:Seoul/Archive2 | 9263684 |
| 6 | Seoul Subway Line 1 | 9074261 |
| 7 | Seoul Foreign School | 7576768 |
| 8 | Seoul Metropolitan Subway | 5946441 |
| 9 | Seoul International School | 4911143 |

Page 97 / 99

과제

1. RSQLite와 DBI를 활용해서 **nycflights13** 데이터를 db Table로 만들고, 제출해 주세요.
 1. 'nycflights13' 패키지를 설치하고 5개 데이터를 확인하세요.
 2. dbConnect 명령으로 SQLite 파일을 sql_[이름].db 으로 생성하세요.
 3. 5개의 데이터를 각각의 이름으로 table을 생성하세요.
 4. 1)~3)의 과정을 모두 코드로 남기고 run_[이름].R로 저장하세요.
 5. run[이름].R과 sql[이름].db 두 개의 파일을 class2assignment 폴더에 저장하세요.
 6. github에서 pull request로 제출해 주세요.
 7. recomen 폴더에 있는 데이터 6개도 같은 과정을 진행하고, 제출은 하지 마세요.
 8. .gitignore를 이용하면 push의 범위에서 제외할 수 있습니다.

과제

1. recomen 폴더에 있는 데이터 6개를 bigquery에 업로드 하고 질의를 실행해 보세요.
 1. "bigrquery" 패키지의 기능으로 작은 용량 5개의 데이터를 업로드해 주세요.
 2. Storage 서비스를 이용해서 **tran.csv** 파일을 bigquery에 테이블로 생성하세요.
 3. query_exec 함수와 "select * from [tran] limit 10" 을 실행하고 결과를 받으세요.
 - [tran]은 각자 해당하는 이름으로 변경하셔야 합니다.
 4. 1)~3)의 과정을 big_[이름].R로 저장하세요.
 5. big_[이름].R 파일을 class2assignment 폴더에 저장하세요.
 6. github에서 pull request로 제출해 주세요.