

8장 과제_2014097056 우성현

2022년 10월 16일 일요일 오후 11:36

8장 레포트(R과 MySQL 연동)

1. R과 R 스튜디오 설치(인터넷을 검색하여 R과 RStudio를 설치함)

2. RMySQL 설치 및 실행

RStudio에서 `install.packages("RMySQL")` 명령어를 실행하면 R과 MySQL 연동 환경이 실행됨.

```
> install.packages("RMySQL")
```

```
> library(RMySQL)
```

3. R에서 아파트 매매 정보 검색

- RStudio에서 "report" 데이터베이스 연결 방법 예

`dbListTables(con)`: 테이블들을 살펴 봄

`dbGetQuery(con, "select * from ApartPrice limit 5")`: 내용을 검색함

```
> library(RMySQL)
> con <- dbConnect(MySQL(), user="root", password="wptkwkd0329-", dbname="report", host="127.0.0.1")
> dbListTables(con)
[1] "apartprice"
> dbGetQuery(con, "select * from ApartPrice limit 5")
  no  ymd      ym year  code  addr1      apt_nm
1  1 2021-01-14 2021-01-01 2021 11110 서울_종로 청운현대
2  2 2021-01-07 2021-01-01 2021 11110 서울_종로 광화문스페이스본
3  3 2021-01-15 2021-01-01 2021 11110 서울_종로 경희궁의아침2단지
4  4 2021-01-15 2021-01-01 2021 11110 서울_종로 경희궁의아침2단지
5  5 2021-01-17 2021-01-01 2021 11110 서울_종로 경희궁의아침3단지

      juso_jibun price con_year area floor  py cnt
1 서울특별시 종로구 청운동 56-45 청운현대 130000      2000   130    2 3300    1
2 서울특별시 종로구 사직동 9-1 광화문스페이스본 150000      2008   145    6 3414    1
3 서울특별시 종로구 내수동 71 경희궁의아침2단지 175000      2004   175    4 3300    1
4 서울특별시 종로구 내수동 71 경희궁의아침2단지 175000      2004   175    4 3300    1
5 서울특별시 종로구 내수동 72 경희궁의아침3단지 180000      2004   123   13 4829    1
```

4. R에서 DB 테이블 조작

질의: aptprice 테이블에서 거래금액(price)가 300,000보다 비싸고, 전용면적(area)가 250보다 큰 아파트이름(apt_nm)과 주소(juso_jibun)을 검색하라.

```
> dbGetQuery(con, "SELECT apt_nm,juso_jibun FROM aptprice WHERE price>300000 AND area>250" )
  apt_nm                                     juso_jibun
1   나인원한남                        서울특별시 용산구 한남동 829 나인원한남
2   다이너스티힐                    서울특별시 서초구 방배동 1-11 다이너스티힐
3   방배롯데캐슬OSCAR                서울특별시 서초구 방배동 1-21 방배롯데캐슬OSCAR
4   청담동마크노빌                    서울특별시 강남구 청담동 89-12 청담동마크노빌
5   상지리츠빌카일룸3차                서울특별시 강남구 청담동 129-16 상지리츠빌카일룸3차
6   상지리츠빌카일룸3차                서울특별시 강남구 청담동 129-16 상지리츠빌카일룸3차
7   상지리츠빌카일룸3차                서울특별시 강남구 청담동 129-16 상지리츠빌카일룸3차
8   청담동마크노빌                    서울특별시 강남구 청담동 89-12 청담동마크노빌
9   청담동마크노빌                    서울특별시 강남구 청담동 89-12 청담동마크노빌
10  PH129                             서울특별시 강남구 청담동 129 PH129
```

질의: aptprice 테이블에서 건축년도(con_year)가 2004년이고 평당매매가(py)가 10,000보다 비싼 아파트이름(apt_nm)과 주소(juso_jibun)을 검색하라.

```
> dbGetQuery(con, "SELECT apt_nm,juso_jibun FROM aptprice WHERE con_year = 2004 AND py>10000" )
  apt_nm                                     juso_jibun
1   아이파크 서울특별시 강남구 삼성동 87 아이파크
2   아이파크 서울특별시 강남구 삼성동 87 아이파크
3   아이파크 서울특별시 강남구 삼성동 87 아이파크
4   아이파크 서울특별시 강남구 삼성동 87 아이파크
5   아이파크 서울특별시 강남구 삼성동 87 아이파크
6   아이파크 서울특별시 강남구 삼성동 87 아이파크
```

5. R에서 데이터 분석하기

질의: aptprice 테이블에서 지역코드(code)별 평당매매가 간의 상관관계를 그래프로 그려본다.

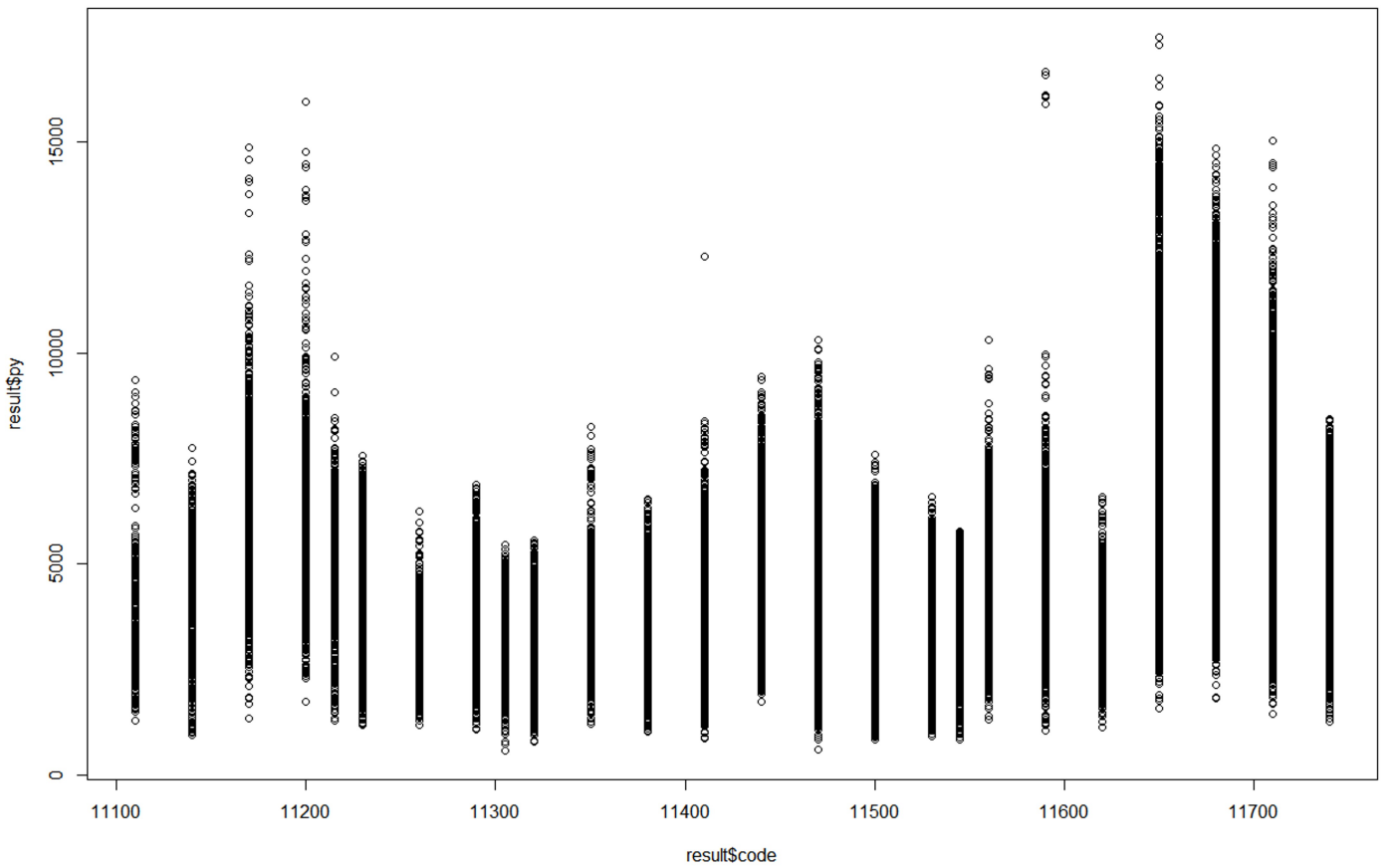
힌트: dbGetQuery 한 내용을 result에 임시 저장

```
result <- dbGetQuery( )
```

result\$code를 x축, result\$py를 y축으로 plot를 이용해 그래프를 그려봄

```
plot( )
```

```
could not run statement: table 'report.aptprice' doesn't exist
> result <- dbGetQuery(con, "select code, py from ApartPrice")
Warning message:
> plot(result$code, result$py)
```



2) GGscatter

```
> ggscatter(result, x = "code", y = "py",
+           add = "reg.line", conf.int = TRUE,
+           cor.coef = TRUE, cor.method = "pearson",
+           xlab = "Region code", ylab = "sale price per sqm")
```

