

06. 데이터 베이스 이용 # src: 7_R/06_데이터베이스이용 참고

1) **sqldf()**; 데이터 프레임의 데이터를 SQL문장을 이용하여 조회 가능하게 해줌**sqldf(x, stringsAsFactors=F)**2) **SQL SELECT 구문** #컬럼명(열 이름)에 . 포함될 경우 반드시 ` (backtick)안에 기재

SELECT [DISTINCT] { * 열이름 [[AS] 열별칭], ... }	#DISTINCT: 중복행 제거
FROM	데이터프레임 이름
JOIN	데이터프레임 이름
ON	조인 조건
WHERE	조건
GROUP BY	그룹화할 칼럼(열) #FACTOR별 집계 가능
HAVING	그룹 조건
ORDER BY	정렬할 칼럼 [[ASC; 오름차순] DESC;내림차순]
LIMIT [m,] n	데이터 조회 개수 지정(m번째 행부터 n개/상위 n개 행) #0부터 시작

(1) 중복행 제거: **DISTINCT** ex) **sqldf('select distinct Species from iris')**(2) 데이터 조회 개수 지정: **LIMIT** ex) **sqldf("select * from iris limit 5")**(3) **sqldf**로 평균, 표준편차 구하기: **avg("컬럼명"), stdev("컬럼명")**#**sqldf('select(avg("컬럼명") as avg, stdev("컬럼명") as sd from data group by 컬럼명(팩터))**3) **RJDBC**; R에서 JDBC를 이용하여 데이터 베이스에 연결하기 위한 패키지(0) 자바 설치 및 환경변수(JAVA_HOME) 설정, 작업디렉토리(getwd())에 **ojdbc6.JAR** 복사#**ojdbc6.JAR (C:\Woraclexe\app\Woracle\product\11.2.0\server\jdbc\lib)**(1) 패키지 설치 및 로드: **install.packages("RJDBC"), library(RJDBC)**(2) JDBC 드라이버 클래스 로드: JDBC객체 생성 **JDBC(driverClass="", classPath="")**① **driverClass="driverClassName";** JDBC 드라이버 클래스 파일명 #**"oracle.jdbc.OracleDriver"**② **classPath="driverClassPath";** JDBC 드라이버 클래스 파일(jar 파일) 위치(경로+파일)(3) 데이터 베이스 연결: **dbConnect(drv, ...)** #접속 URL, 사용자ID, PW(4) 테이블 데이터 조회 #**con: dbConnect()**함수가 반환한 **JDBCConnection** 드라이버 객체① **dbGetQuery(con, "SELECT문", ...);** 테이블 데이터 조회 #**SELECT문만 사용**, 데이터프레임 반환② **dbReadTable(con, "name", ...);** 테이블 전체 데이터 조회 #**"name": 데이터 베이스 테이블명**(5) 테이블 데이터 수정: **dbSendUpdate(con, "INSERT/UPDATE/DELETE문", ...), dbExecute()**(6) 데이터 베이스 연결 종료: **dbDisconnect(con)**