

[데이터 전처리-III]

8. 데이터프레임 이름 생략

(1) **with()** / **within()**

; 이 함수들을 데이터 프레임의 필드를 데이터이름없이 접근할 수 있다.

; with절은 결과값만을 반환, within절은 결과값이 포함된 데이터 프레임셋 전체를 반환

(2) **attach()** / **detach()**; 데이터이름 생략해도 되도록 설정 / attach기능 해제

ex) iris 데이터에 attach와 detach설정하기

attach(iris)

Sepal.Length # attach 설정 후엔 iris\$Sepal.Length라고 입력하지 않아도 실행됨

detach(iris) # attach 기능 해제

9. 데이터 집계

(1) **table()**: 데이터 개수

ex) iris데이터의 종별 iris 수

table(iris\$Species)

(2) **aggregate**(대상자료, **by=list**(그룹화할 요소 목록), 함수)

: 데이터를 하위집합으로 분류하고 요약 통계 계산

ex) iris의 꽃받침 길이와 넓이, 꽃잎의 길이와 넓이의 종별 평균

aggregate(iris[,1:4], by=list(iris[,5]), mean) # 1~4열: 꽃잎, 꽃받침 정보 / 5열: 종류

10. 조건으로 색인 찾기 **which()**

ex) iris 데이터에서 'setosa'종만 추출

which(iris\$Species == 'setosa') # 색인만 추출

iris[**which**(iris\$Species == 'setosa'),] # 데이터 추출