

데이터 형태

범주형 데이터 (categorical data)

- 관측 결과가 몇개의 범주 또는 항목의 형태로 나타나는 자료

• (예) 성별(남,여), 선호도(좋다,보통,싫다), 혈액형(A,B,O)

- 범주형 데이터를 수치형 자료처럼 표현 가능하다.

• (예) 남=0, 좋다=0, 싫다=1 보통=2

- 수치형 자료처럼 표현되어있는 범주형 자료를 잘 구분해야 한다.

• (남=0) (여=1) 하게되면 여자가 '1' 이라고 묵히에 있을것 아니기때문

- 순위형 자료 명목형 자료로 나뉜다.

• 순위형 자료 (ordinal data)
- 범주간에 순서가 의미가 있는 자료
(예) (좋다 보통 싫다)처럼 범주간 순서가있는것
• 명목형 자료 (nominal data)
- 범주간에 순서가 의미가 없는 자료
(예) (혈액형, 성별)

수치형 데이터 numerical data

- 고난도 값이 수치로 측정되는 자료

(예) 키, 몸무게, 시험 성적

- 연속형과 이산형 자료로 구분된다.

• 연속형 자료 continuous data
- 키, 몸무게 같은 값이 연속적인 자료
• 이산형 자료 discrete data
- 자동차와 사고 건수와 같이 값이 셀수있는 자료