



## 빅데이터 예측분석 9

서강대학교 정보통신 대학원 정 화민 교수 (MIS Ph.D.)



# 자료분류/ 표현

## ○ 수량화 여부에 따른 구분

### - 질적 자료 (Qualitative Data) / 범주적 자료 (Categorical Data)

. 색깔,성,지역,직업,설문조사결과 등 범주화/식별성 만 가능

\* 자료의 식별 방법 ➡ 식별,코드,상품코드(바코드,QR코드,RFID 등) 참조

\* 컴퓨터 내 문자 표현 (연산 불가) ➡ 알파뉴메릭 코드(ASCII,유니코드 등)

### - 양적 자료 (Quantitative Data)

. 나이,소득,매출액 등 수량화 가능 (수량적 자료)

.. 계수치(이산적) : 셀 수 있는 정수 값 (例, 생산개수,불량수 등)

.. 계량치(연속적) : 연속 값 (例, 길이,무게,인장강도,온도,시간 등) .

# 자료 분류/표현

---

## ○ 속성에 따른 구분

### - 질적 자료 / 범주 자료

- . 명목 자료(Nominal Data) : 어떤 속성을 분류하기 위해 수치 부여 (예, 남 1, 여 0 등)
- . 서열 자료(Ordinal Data): 상대적인 크기 비교 (예, 사회계급, 선호순위 등)
  - \* 사칙연산 의미 없음

### - 양적 자료

- . 등간 자료/구간 자료(Interval Data)
  - .. 간격이 균등한 수치 부여 (예, 온도, IQ 등)
- . 비율 자료(Ratio Data)
  - .. 균등 간격에 절대 영점 있음 (예, 나이, 소득, 무게 등)
  - \* 사칙연산 가능

## 자료 분류/표현

	양적(계량적) 자료	질적자료
목적	현상간의 관계를 계량화 현상간 관계의 일반화	현상에 대한 깊은 이해
표본	대표성 있는 큰 표본	대표성 없는 작은 표본
자료수집	구조화된 자료수집	구조화되지 않은 수집
자료분석	통계분석	비정형 통계분석

양적(계량적)자료: 현상을 계량적인 의미를 가지는 부호, 즉 수로 측정한 자료

질적자료: 비계량적인 부호(즉 언어)로 측정한 자료

## 자료 분류/표현

척도	제공정보	예	대표값	분석예
명목척도	카테고리 구분	주민등록번호, 성별, 직업, 근무부서	최빈값	빈도분석, 교 차분석, 비모수통계
서열척도	카테고리 구분, 순위	선호순위, 등 수, 사회계층	중앙값	서열상관분석, 비모수통계
등간척도	카테고리 구분, 순위, 간격비교	선호도, 태도, IQ, 온도계, 직무만족, 이 직의사	평균	모수통계
비율척도	카테고리 구분, 순위, 간격비교, 절대크기 비교	매출액, 이익, 나이, 몸무게, 연봉, 근무연 수	평균	모수통계

박사학위논문

기업의 빅데이터 활용 유형별  
적용 방안 및 효과 분석

A Case Study on Application Methods and  
Effects based on the Type of Big Data  
at Corporations

2014년 8월

한신대학교 대학원

정보통신학과

이재성

참조논문  
카페 확인