

제 1 장 통계학의 기본개념

1. 통계학이란 무엇인가?

통계학(statistics)이란, 불확실한 상황에서 해결하고자 하는 문제를 대상으로, 주어진 자료를 수집·정리하며 체계적으로 분석하는 학문

1.1 불확실성 Uncertainty

관찰이 부족하거나 확률적 상황일 때

예 1. 전체 5명의 성적 {90, 90, 100, 80, 90}

첫 3명의 성적만으로 전체 5명의 평균을 추정할 때

예 2. 주사위 던질 때 나올 값

1.2 통계학 어원

Statistik(German), 1749, Gottfried Achenwall, the analysis of data about the state

-18세기: systematic collection of demographic and economic data by states

-19세기: the discipline concerned with the collection, summary and analysis of data

2. 통계학의 기본용어

2.1 자료의 성격

자료(Data), 정보(Information)와 통계학(Statistics)

2.1.1 자료[Data]

자료란 사람, 물체, 조건 또는 상황을 묘사하는 상징으로서 수량, 시간, 금액, 비율, 이름, 장소 등을 포함하는 기본적인 사실들의 집합이다.

2.1.2 정보[Information]

정보는 자료를 의미있는 형태로 전환시킨 것

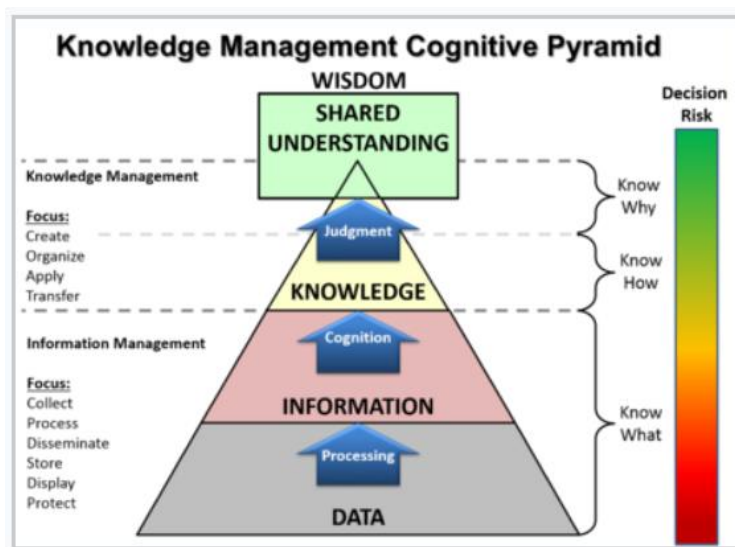
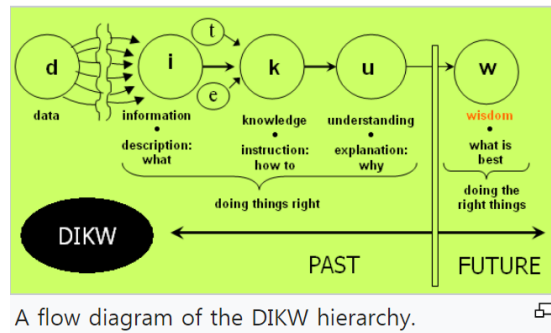
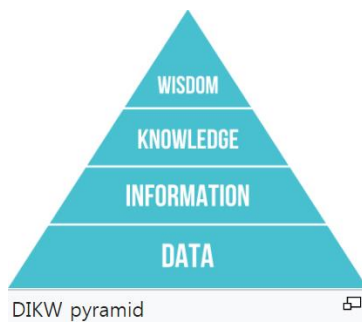
통계학은 자료를 가치있게 변화시켜 정보를 제공하는 학문

2.1.3 자료와 정보

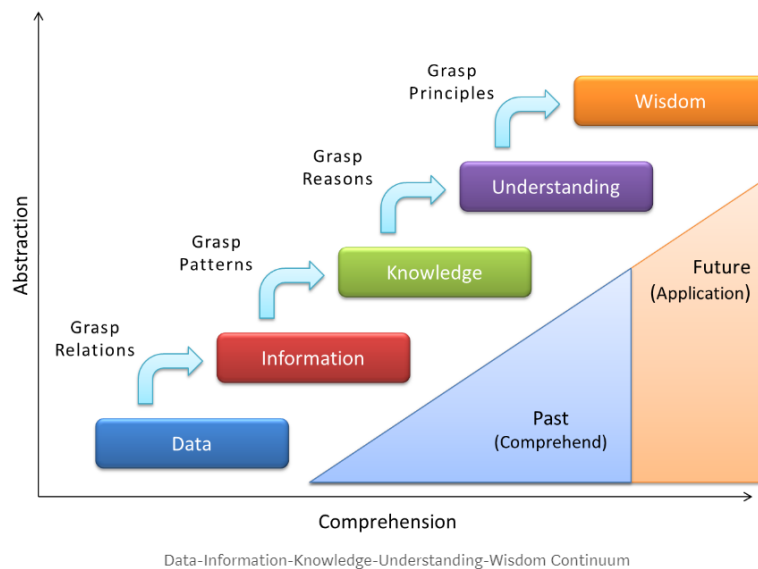
data는 그 자체만으로는 의미를 전달하지 못한다.

data가 information이 되기 위해서는 해석되어야 하고, 의미를 지녀야만 한다.

2.1.4 [심화학습] Data-Information-Knowledge-Wisdom Pyramid (Wikipedia)



DIKUW Continuum



2.1.4.1 관련 동영상

What is DIKW? - <https://www.youtube.com/watch?v=h9gYk66yz-0>

Data Information and Knowledge - <https://www.youtube.com/watch?v=QsP5WGv0aQc>

2.2 모집단과 표본

모집단: population, 연구자의 조사 대상이 되는 전체 측정치의 집합
(예: 특정 회사 제품 전체)

표본: sample, 모집단의 일부 측정치 (예: 전체 제품 중 발췌된 일부)

2.3 기술통계학과 추리통계학

2.3.1 기술통계학(descriptive statistics)

통계자료를 적절하게 그림, 도표 또는 수치로 요약하고 그 특성을 기술하는 방법론
통계적 자료를 정리하고 요약하는 등, 자료특성의 계산방법과 관련된 통계학

2.3.2 추리통계학(inferential statistics)

모집단에서 뽑은 표본을 통하여 모집단의 특성을 예측 또는 추측하는 방법론
모집단에서 뽑은 표본을 분석하여 이를 기초로 모집단의 특성을 규명하는 통계학

2.3.2.1 통계적 추론(statistical inference)

표본을 사용하여 추정하는 과정(the process of drawing conclusion from data subject to random variation)

2.3.2.1.1 통계적 추론의 예

특정 후보에 대한 유권자 전체의 지지율을 알고 싶어 함. 전체 유권자 중 천 명을 상대로 특정 후보의 지지율을 파악. 천 명의 지지율을 토대로 전체 유권자의 지지율을 추정함

모집단: 유권자 전체

표본: (유권자 전체 중) 천 명

모집단 특성: 특정 후보에 대한 유권자 전체의 지지율

통계적 추론: 특정 후보에 대한 천 명 지지율을 토대로 전체 유권자 지지율 추정