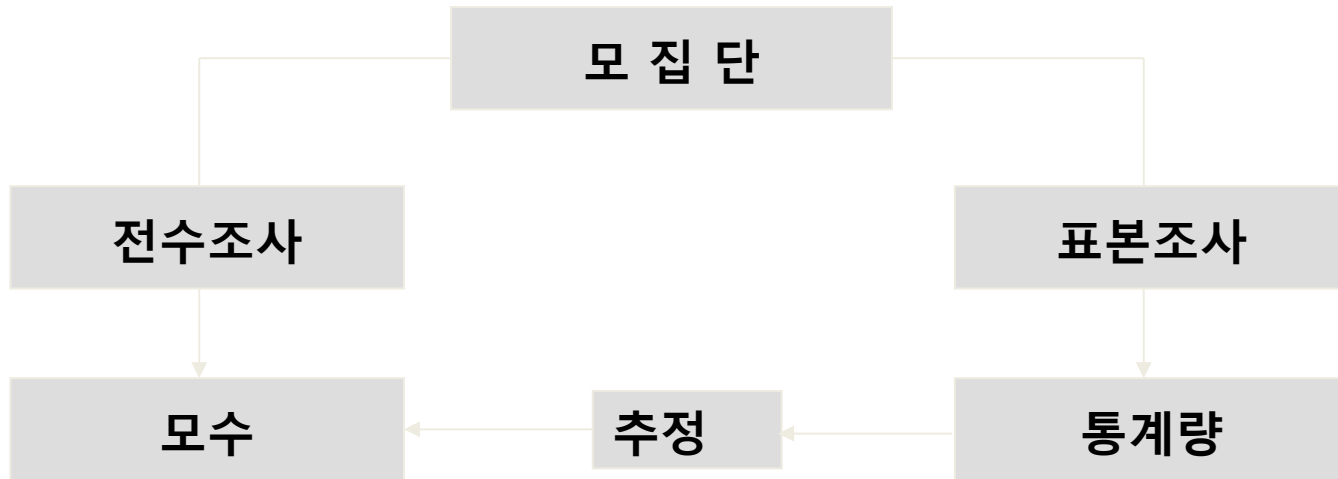

5. 연구대상자 선정 (표본추출)

전수조사~표본조사



μ : 모평균
 σ : 모표준편차
 σ^2 : 모분산



$\bar{x}$: 표본평균
 S : 표본표준편차
 S^2 : 표본분산

- 전수조사(센서스 ; *census*, complete numeration)
 - 모집단에 속하는 모든 구성원들을 대상으로 조사하는 것.
ex) 인구 주택 총조사
- 표본조사 (sample survey)
 - 관심의 대상이 되는 전체 집단 중 일부의 부분집단을 선택한 후 그 일부 집단을 조사하여 얻은 데이터를 이용해서 전체 집단의 특성을 추정하는 통계조사.
 - 모집단을 대표할 수 있어야 한다.
 - 표본으로부터 모집단의 하나 혹은 그 이상의 특성들을 추정
 - 명확한 모집단의 정의
 - 표본의 크기는 모집단의 크기에 따라 결정

전수조사와 표본조사

구분	전수조사	표본조사
예산 금액	많은 경우	적은 경우
허용되는 시간	긴 경우	짧은 경우
모집단의 수	작은 경우	큰 경우
특성에 대한 분포	분산이 큰 경우	분산이 작은 경우
표본오차의 문제	많은 경우	적은 경우
비표본오차의 문제	적은 경우	많은 경우
측정의 형태	비파괴성 측정	파괴성 측정

표본조사의 장_단점

• 장점

1. 경제성
2. 신속성
3. 비표본오차, non-sampling error) 감소
4. 정확성의 확보 - 소수의 잘 훈련된 조사원
 - 면접, 관찰, 측정이 정확하지 못한 경우
 - 조사표 기입사항의 검토가 불충분하거나 부호화, 정리상의 오차
 - 추출대상(frame)이 불완전하여 조사단위의 일부가 누락 또는 중복된 경우
5. 파괴적인 조사에의 적용가능
6. 전수조사가 불가능한 경우
 - 새로 개발한 의약품의 효과
7. 다량의 정보 확보가능

• 단점

- 대표성 있는 표본선정의 어려움
- 모집단 자체가 작은 경우에는 무의미
- 모집단의 세부적인 특성을 알기 어려움

표본추출 관련 용어 1

- 모집단(population)
 - 우리가 관심을 갖는 사람, 개체, 생물 전체
- 개체(unit): 모집단의 구성원
 - 추출단위 : 표본으로 추출되는 단위
 - 기본단위 $\leq$ 추출단위 : 개인 or 가구를 표본 ?
- 표본(sample)
 - 모집단의 일부로 모집단의 정보를 모두 갖추고 있어야 좋은 표본이다.
- 표본추출(*sampling*)
 - 특성을 알고자 하는 어떤 대상(전체 : 모집단)의 일부분을 선택하는 것으로 일부분의 특성을 이해함으로써 전체의 특성을 추정할 수 있다는 가정하에서 표본추출이 이루어진다.
- 추출틀(sampling frame) : 명부, 약도, 카드
 - 모집단의 모든 추출단위를 포함, 누락이나 중복 제거, 명확하고 쉽게 식별이 가능
- 변수(variable)
 - 표본 조사에서 측정하게 되는 개체들의 특성, 설문 조사에서는 이것이 설문 항목이 된다.
- 모수(parameter), 통계량(statistic)
 - 추정량(estimator):모수를 추정하기 위한 통계량
 - 추정값(estimate):하나의 표본에 의해 얻는 추정량의 실현값

표본추출 관련 용어 2

- 표적집단
 - 최종적으로 그 연구 결과를 적용하고자 하는 궁극적인 집단
- 모집단
 - 그 규모를 파악할 수 있는 인구집단으로서, 연구자가 접근 가능하고, 연구결과를 일반화하고자 하는 인구집단
- 표집집단
 - 모집단을 대표할 수 있는 표본으로서 적절한 표본추출(표집)과정을 거쳐 선정된 집단
- 적격집단
 - 표집집단 중 연구 목적에 부합되는 적절한(eligible) 대상을 의미

표적집단

65세 이상 한국인

모집단

익산시 거주 65세 이상 노인(n=16,910)

표집집단

모집단에서 추출된 65세 이상 노인
(n=1,600)

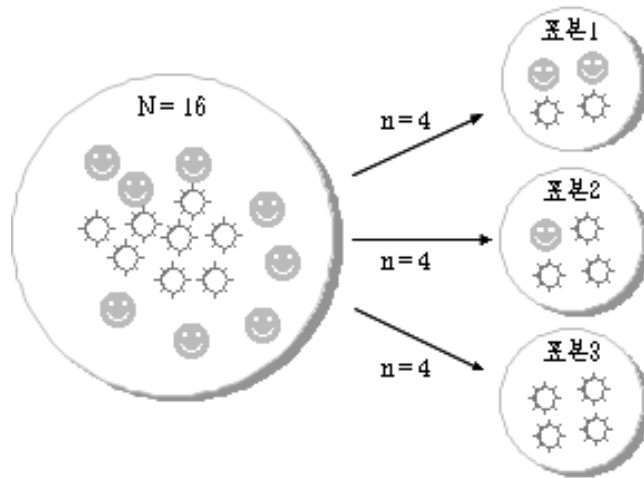
적격집단

표집집단 중 조사 당시 생존, 주소 정확,
해당 거주지에 계속 거주하였던 사람
(n=1,330)

참여집단

적격집단 중 실제 조사에 참여한 사람
(n=940)

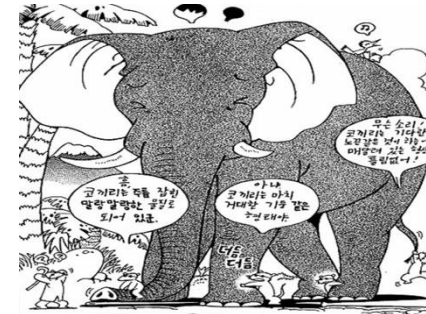
표본추출 관련 용어 3



- 표본오차
 - 표본 1은 모집단과 같은 특성으로 구성되어 있지만, 표본 2와 표본 3은 모집단과는 다른 특성의 표본으로 구성되어 있다.
 - 이때 표본 2와 표본 3을 이용하여 조사하는 경우 모집단을 조사하였을 때 얻을 수 있는 결과와는 다른 결과가 나타날 수 있으며, 이것을 표본오차(sampling error)라고 한다
- 비표본오차
 - ✓ 표본오차를 제외한 나머지 모든 오차
 - ✓ 조사기획단계, 조사원, 표본프레임 등에 의한 오차
 - ✓ 전수조사(Census)에서는 비표본오차만 존재

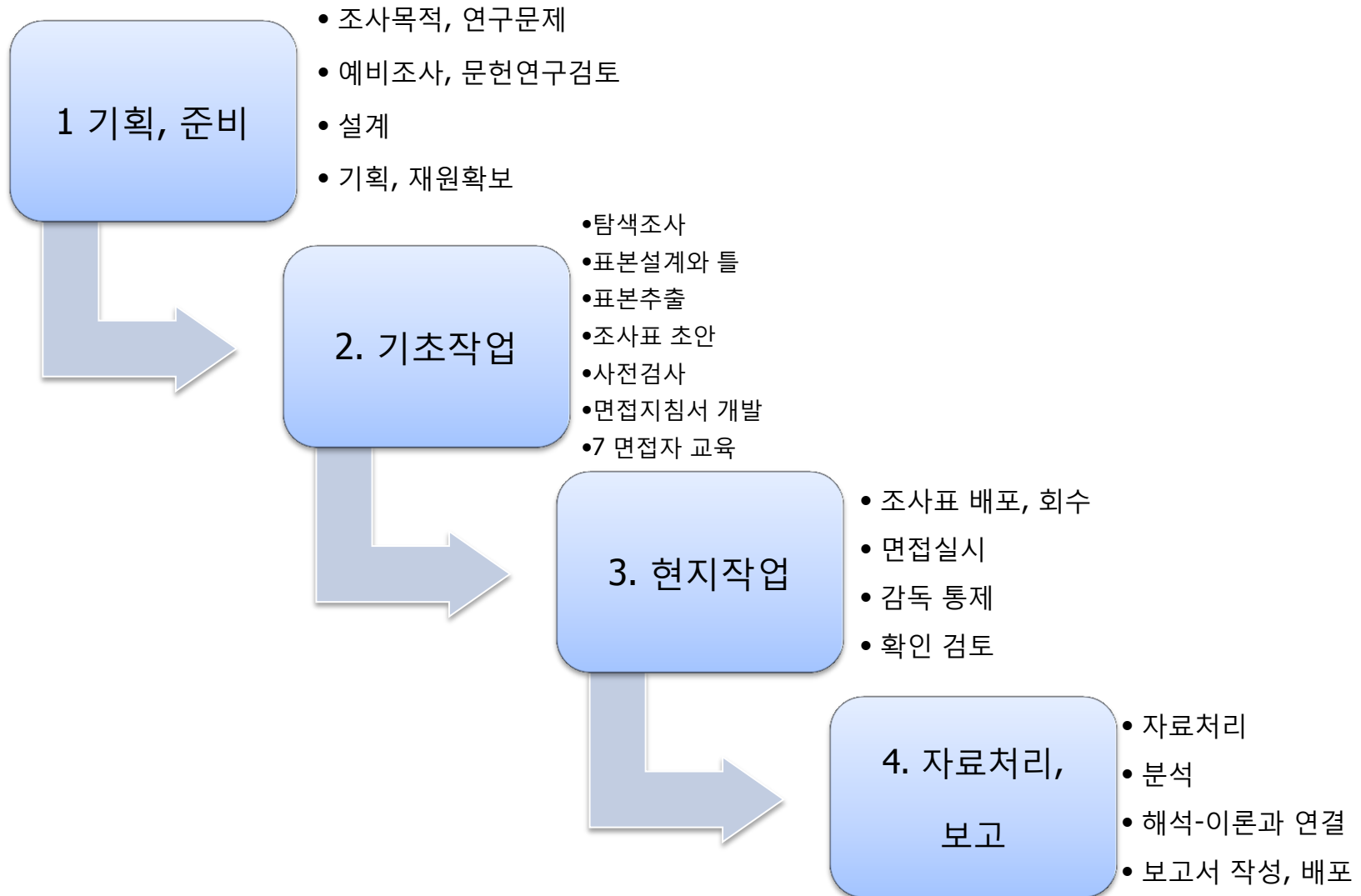
표본추출의 오류

- 1935년 '다이제스트'지는 당시 대통령이던 프랭클린 루즈벨트 와 앨 프리드 랜던 상원의원이 1936년 대통령 선거에서 랜던 상원의원이 승리할 것이라는 추측기사.
- 유권자 1,000 만명에게 누구를 찍을 것인지 설문조사
- 그 중 230만개 회수
- 3:2로 랜던 승리 예측

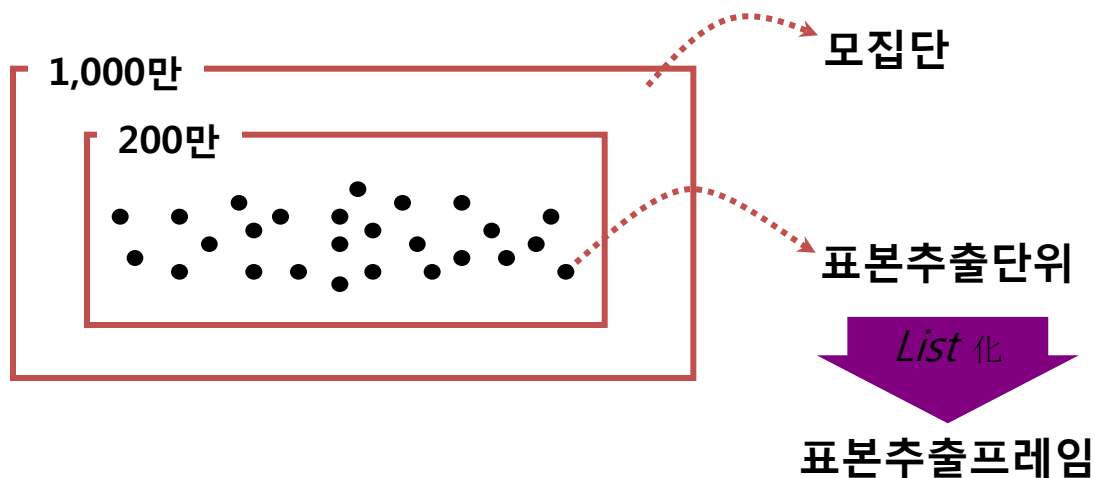


표본추출 방법

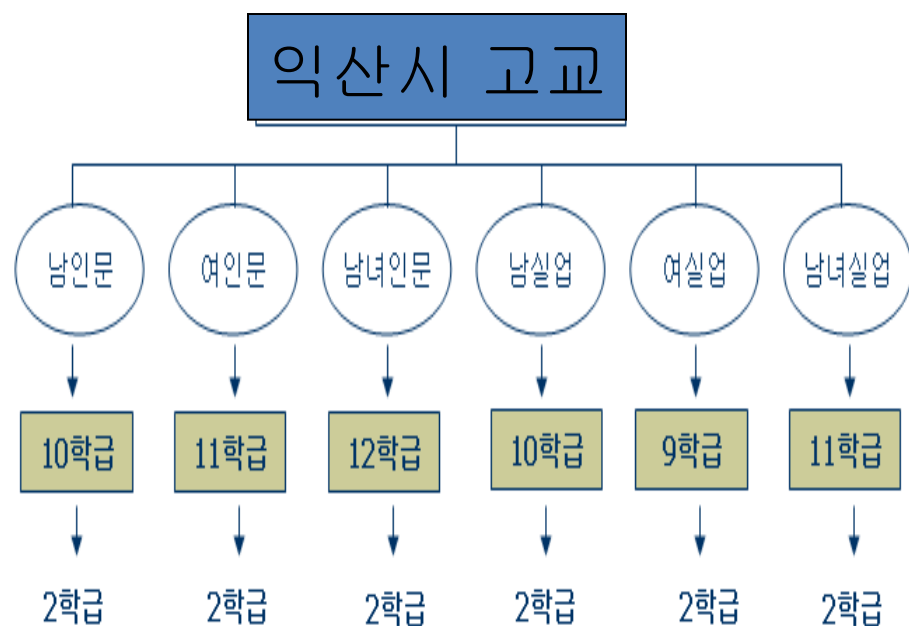
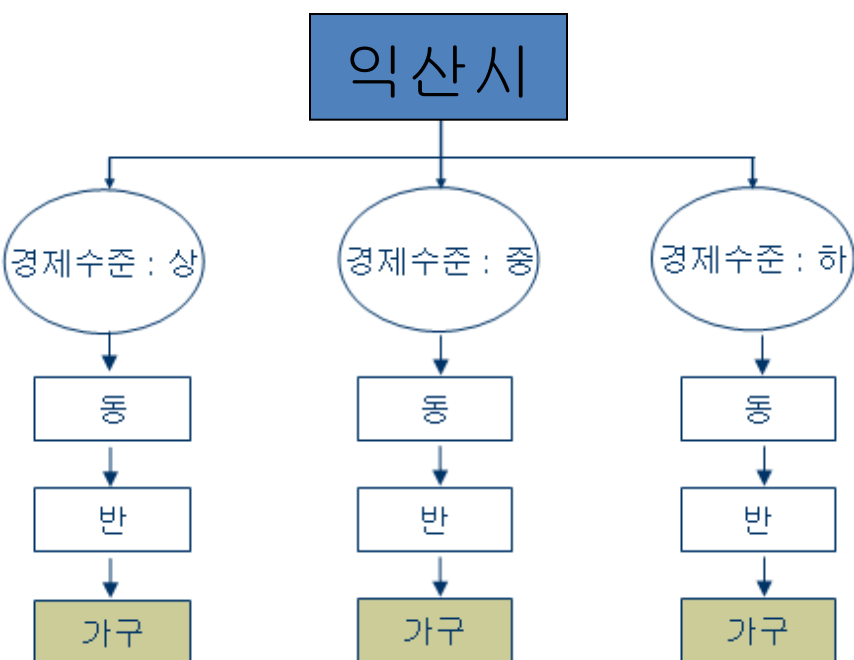
Survey Research Process



- 구성원(*element*) : 자료 혹은 정보를 수집하는 대상.
- 표본추출단위(*sampling unit*) : 표본추출과정에서 선택의 대상이 될 수 있는 표적모집단의 구성원.
 - 단순/일단계 표본추출(*simple/single-stage sample*)의 경우 : 표본단위는 모집단 구성원과 동일.
 - 조사대상 지역이 넓은 경우 표적모집단 구성원들 중 일부는 표본단위에서 제외될 수 있다. 이는 비용과 시간 경제적인 이유도 있으며 표본에 의한 편견(*bias*)을 방지할 수도 있기 때문이다. 이 경우 표본단위 수는 표적모집단 구성원들의 수보다 적다.
- 표본추출프레임(*sampling frame*) : 표본추출을 실행하기 위해서 표본단위들로 구성된 표본프레임이 마련되어야 하며, 표본추출프레임이란 모든 표본단위들의 리스트(*list*)를 말한다.



표본추출단위의 예



표본추출의 단계

단계 1	모집단(<i>population</i>)의 결정
------	------------------------------

단계 2	자료수집방법의 결정
------	------------

기술조사/인과관계조사를 위한 자료수집방법 : 서베이법(대인인터뷰 – 방문/전화인터뷰/우편서베이 등), 관찰법, 실험법 ← 어떤 방법으로 자료를 수집할 것인가 결정

단계 3	표본추출프레임의 결정
------	-------------

자료수집방법과 관련되며, 표적모집단과 일치하는 것이 이상적이나 샘플링 디자인을 잘못하면 표본추출프레임 오류(*sampling frame error*)가 발생할 수 있다.

표본추출프레임 오류가 있는 경우의 표본조사를 하여 해석한다면 현상에 대한 잘못된 이해(*mislead*)를 초래하나 현실적으로 비용, 시간 경제적인 이유로 표본추출프레임으로 표적모집단보다 작게 구성하는 경우가 드물지 않으며, 일부 표본추출프레임을 제외하면 표적모집단보다 작아지지만 비용과 시간을 크게 줄일 수 있다는 측면에서 표본추출프레임 오류는 양해될 수 있다.

단계 4

표본추출방법의 결정

- 확률표본추출(*probability sampling*)과 비확률표본추출(*nonprobability sampling*)의 두 가지 방법이 있으며 확률표본추출을 하는 것이 모집단의 특성을 측정하는데 보다 바람직하나 비확률표본추출은 나름대로 장점을 가지며 현실적으로 많이 사용된다.

단계 5

표본크기 결정

- 고려사항
 - 표본크기 결정을 위해 추정치(*estimates*)가 얼마나 정확해야 하는지, 얼마나 많은 시간과 비용을 사용할 수 있는지를 고려해야 한다 : 다른 조건이 같다면 표본이 클수록 추정치가 정확하며 시간과 비용이 많이 소요되고, 모집단의 특성의 분산이 클수록 (모집단 구성원이 특성이 다양할수록) 보다 큰 표본이 요구된다.
 - 얼마나 많은 구성원들을 대상으로 조사할 것인가 : 현실적으로 소비자들을 대상으로 설문조사를 실시할 때 조사대상자로 결정되더라도 응답을 거부하는 경우가 많다. 또한 응답을 하더라도 여러 가지 이유로 유효한 자료로 사용하지 못하는 경우도 흔하므로 이러한 점들을 고려하여 접촉할 표본의 크기를 결정해야 한다.

단계 6 표본추출 실행계획 수립

- 조사자 : 조사대상(자)에게 어떻게 접촉할 것인지를 결정.

단계 7 표본추출의 실행

- 표본추출프레임이 구체적으로 설정된 경우 표본추출 계획수립자가 표본추출을 할 수 있지만, *Mall-Intercept* 조사처럼 표본추출프레임 없이 조사가 실시되는 경우는 면접원이 지침에 따라 표본을 추출하고 조사하게 됨.

확률표본추출법과 비확률표본추출법의 비교

확률표본추출	비확률표본추출
단순무작위표본	편의표본
체계적 표본	판단표본
층화표본	할당표본
❖ 비례적	
❖ 불비례적	
군집표본	눈덩이표본
❖ 지역표본	

여러 가지 방법이 발생하는 이유는 적은 시간에 적은 비용으로 효과적인 결과를 얻기 위해 각 상황에 적당한 최적방법을 사용하기 위해서 이다.

모집단에 대해 보다 대표성이 높다.

비교기준	확률표본추출	비확률표본추출
표본의 모집단 대표성	높음	낮음
표본추출 오류계산	가능	불가능
추계통계기법 적용	가능	불가능
비용	높음	낮음
표본추출기법	높은 수준 요구됨	높은 수준 요구되지 않음

표본추출방법의 선택

- 표본오차는 계산 가능하며 최소화할 수 있는가?
- 추정량과 그에 대한 분산은 직접 계산 가능한가?
- 가용예산과 인원의 제약 아래에서 요구정도를 달성할 수 있는가?
- 표본조사를 실행하는 과정은 용이한가?
- 유사한 조사방법을 과거에 실행한 적은 있는가?