사회통계연습

김현우, PhD¹

¹ 충북대학교 사회학과 부교수



진행 순서

- 자료구조 재배열
- ② 자료관리 종합연습





분석목적에 따라 자료를 뒤틀어 다른 꼴로 바꾸어야 할 때가 있다.

- 분석의 목적에 따라 긴 꼴(long form)과 넓은 꼴(wide form)로 자료구조를 재배열 (reshaping)해야 하는 상황이 나온다.
- 자료의 본질은 변하지 않는다. 그저 형태가 바뀌는 것에 지나지 않는다.



- 아래 예제는 넓은 꼴에서 긴 꼴로 자료를 변화할 때 어떻게 달라지는지 보여준다.
- (Weight과 BP 대신) Key와 Value가 새로 생겼다. 그리고 기존의 구체적인 값들은 Value 안에 들어갔다.

				Name	Key	value
Name	Weight	BP				
				John	Weight	150
John	150	120	→	John	BP	120
Smith	170	130		Smith	Weight	170
				Smith	BP	130



• 아래의 또다른 예제는 긴 꼴에서 넓은 꼴로 자료를 변환하는 상황을 보여준다.

country	year	metric
Х	1960	10
Х	1970	13
Х	2010	15
у	1960	20
у	1970	23
у	2010	25
Z	1960	30
Z	1970	33
Z	2010	35

country	yr1960	yr1970	yr2010	
x	10	13	15	
у	20	23	25	
z	30	33	35	

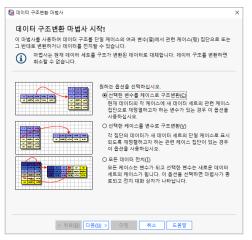


대체 왜 이런 재배열이 필요할까?

- 그 상황을 직접 부딪쳐보면 그 필요성을 이해할 수 있다.
- 만약 국가별로 metric을 종속변수로, year를 독립변수로 하는 회귀분석을 수행한다면 자료가 어떤 꼴일 때 더 편리할까?
- 만약 국가별로 특정 해의 metric 평균을 구하고 싶다면 자료가 어떤 꼴일 때 더 편리할까?
- 문제는 적절한 배열의 형태를 우선 파악하고, 이를 제대로 수행하기가 생각보다 까다롭다는 점이다.
- (자료 결합과 마찬가지로) 혼동스럽기 때문에 상당한 경험과 연습을 필요로 한다.



- SPSS의 [데이터]-[구조변환]에서 볼 수 있는 예시 그림을 자세히 들여다보자.
 - (1) '선택한 변수를 케이스로 구조변환'이 넓은 꼴에서 긴 꼴로 재배열한다.
 - (2) '선택한 케이스를 변수로 구조변환' 이 긴 꼴에서 넓은 꼴로 재배열한다.





연습 1. BODYFAT.SAV 자료의 배열을 파악하고 다른 꼴로 재배열하시오.



연습 2. 재배열된 BODYFAT.SAV 자료를 다시 원래 형태로 복구하시오.



- SPSS에서 넓은 꼴에서 긴 꼴로 재배열하였다면 지수1과 trans1 같은 변수들이 나타난다.
- 여기서 지수1은 사실 Key이고, trans1은 사실 Value임을 알 수 있다(Why?).

				Name	Key	value
Name	Weight	BP				
				John	Weight	150
John	150	120	\longrightarrow	John	BP	120
Smith	170	130		Smith	Weight	170
				Smith	BP	130



연습 $3. \text{ Y1_STD_EDU.SAV}$ 자료의 $\frac{20}{2}$ 을 파악하고 다른 형태로 재배열하시오.



연습 4. 변환된 Y1_STD_EDU.SAV 자료를 다시 원래 형태의 꼴로 복구하시오.



자료관리 종합연습





자료관리 종합연습

연습 5. Y1_STD_EDU.SAV 자료와 Y2_STD_EDU.SAV 자료에서 각각 사회 자기(social self)의 총점을 계산하시오. 전자에 후자를 결합 (merging)한 뒤, 두 사회 자기 변수들의 연관성을 시각화하시오(결측치 처리에 주의할 것).



자료관리 종합연습

연습 6. PROD.SAV 자료에는 1970년부터 1986년까지 미국 48개 주 (state)에서 공공자본스톡(public capital stock)과 민간자본스톡(private capital stock)이 담겨 있다. 이 자료의 관찰 단위(unit of observations)가 현재 무엇인지 판단하고, 이를 주-연도(state-year) 단위로 변환하시오.

