Ver. 1.1

따라하면 끝나는 통계분석

한울통계컨설팅



본 책자는 양적연구로 학위논문이나 학회지를 준비하시려는 분들이 고민하는 통계 분석에 대해서 조금이나마 도움을 드리고자 작성하게 되었으며, 통계에 대해서 아 무런 지식이 없는 분들이라도 데이터 코딩과 같은 기초부터 초급통계, 중급통계까 지 책자를 참고하시면 최대한 쉽게 통계분석을 하실 수 있을 것입니다.

본 책자의 내용은 학위논문이나 학회지에 들어가는 통계 중 가장 많이 활용되고 있는 빈도분석, 기술통계, 교차분석, 차이분석(t-test, ANOVA), 요인분석, 신뢰도 분석을 기본으로 상관분석, 다중회귀분석 및 더미를 활용한 다중회귀분석을 초급통 계에서 다루었으며, 더욱 나아가 위계적 회귀분석, 위계적 회귀분석을 활용한 조절 효과분석, 회귀분석을 활용한 매개효과분석, 이항로직스틱 회귀분석, 공분산분석, 반 복측정분산분석, 정규성 검정, 비모수 통계와 같은 중급통계를 더 추가하여 다루었 습니다.

스스로 통계분석을 공부하시는데 필요한 내용을 10년 이상 통계를 분석해 온 실 무자 입장에서 최대한 쉽게 풀어서 작성하였으며, 다양한 통계서적를 참고하였습니 다. 그렇다고 제가 완벽하다고는 말씀 드리지 못합니다.

본 책자를 참고하시다가 막히는 부분이나 수정해야 해야 할 내용이 발견된다면 아 리 연락처로 연락주세요. 이해가 가지 않는 부분에 대해선 친절히 알려드리고 수정 해야 할 내용이 있다면 내용파악 후 수정하도록 하겠습니다. 또한 이런한 피드백을 통해 더욱 쉽게 따라할 수 있도록 책자를 업그레이드 하여서 무료로 배포하도록 하 겠습니다.

2018년 제작 : 한울통계컨설팅



※ 본 책자는 필요한 분들에 한하여 무료배포는 가능하나 수정 배포는 불가합니다.

※ 목 차 ※

1. 데이터 코딩 방법 4 2. 엑셀과 한글 데이터 SPSS로 옮기기 6 1) 엑셀데이터 SPSS로 옮기기 6 2) 한글 또는 메모장 데이터 SPSS로 옮기기 8 3. 역코딩 13 4. 요인 만들기 15 5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팀 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 28 6. 분산분석(ANOVA) 32 7. 요인분석 44 10. 회귀분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 1) Enter(입력) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 III. 중급 통계 분석 59 1 의게직 다주회귀분석 59
2. 엑셀과 한글 데이터 SPSS로 옮기기 6 1) 엑센데이터 SPSS로 옮기기 6 2) 한글 또는 메모장 데이터 SPSS로 옮기기 8 3. 역코딩 13 4. 요인 만들기 15 5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 26 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 1) Enter(입력) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 II. 중급 통계 분석 51
1) 액셀테이터 SPSS로 옮기기 6 2) 한글 또는 메모장 데이터 SPSS로 옮기기 8 3. 역코딩 13 4. 요인 만들기 15 5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 1. 인구학적 특성에 따른 빈도분석의 팁 21 1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 교차분석 24 독립 t-test 25 다양 다-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 40 회귀분석 41 10. 회귀분석 42 9. 상관분석 44 10. 회귀분석 45 다 (입력) 방식 다중회귀분석 46 1) Enter(입력) 방식 다중회귀분석 47 48 48 11. 더미를 활용한 다중회귀분석 44 51
2) 한글 또는 메모장 데이터 SPSS로 옮기기 8 3. 역코딩 13 4. 요인 만들기 15 5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 1) Enter(입력) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 II. 중급 통계 분석 59 1 의계적 다중회귀분석 59
3. 역코딩 13 4. 요인 만들기 15 5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팀 21 2. SPSS 아웃폿을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 59 1< 위계적 다준회귀분석
4. 요인 만들기 15 5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 59 1< 위계적 다주희귀분석
5. 연속형 데이터를 범주형 데이터로 바꾸기 17 6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팀 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 44 10. 회귀분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 III. 중급 통계 분석 59 1< 위계적 다중희귀분석
6. 이상값 찾기 20 II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 III. 중급 통계 분석 59 1 위계적 다준회귀분석 59
II. 초급 통계 분석 21 1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 51 III. 중급 통계 분석 59 1. 외계적 다중회귀분석 59
1. 인구학적 특성에 따른 빈도분석의 팁 21 2. SPSS 아웃풋을 엑셀로 보내는 방법 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1. 외계적 다중회귀분석 59
1. 친구덕가덕 8 때까근 친고한다더 1 21 2. SPSS 아웃풋을 엑셀로 보내는 방법······23 23 3. 교차분석 26 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1< 위계적 다중히귀분석
2. 51 55 가 () (2) (2) (2) (2) (2) (3) (
3. 표시분곡 20 4. 독립 t-test 28 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 III. 중급 통계 분석 59 1 위계적 다중회귀분석 59
1. 여급 t test 20 5. 대응 t-test 30 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
0. 대장 다 test 50 6. 분산분석(ANOVA) 32 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
0. 한 만한 (untovin) 52 7. 요인분석 38 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
1. 보인 1 50 8. 신뢰도 분석 42 9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
9. 상관분석 44 10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
10. 회귀분석 46 1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
1) Enter(입력) 방식 다중회귀분석 46 2) Stepwise(단계선택) 방식 다중회귀분석 48 11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1<위계적 다중회귀분석
2) Stepwise(단계선택) 방식 다중회귀분석 ····································
11. 더미를 활용한 다중회귀분석 51 Ⅲ. 중급 통계 분석 59 1 위계적 다중회귀분석 59
Ⅲ. 중급 통계 분석
비. 중집 중계 군구 ···································
9 이게저 친구부서 약 하 위 친 · 조정 중 과부서 ···································
2. 귀계적 외위군적을 철중한 조절효과군적
3. 외위문극을 실당한 배개요가 군국 ···································
4. 이용도시드릭 외대군국 74 5 고부사부서
0. 8 한 한 한 국 ·····························
0, 근ㄱㄱ o ᆫᆫᆫ 7 저규서 거저
·· 이미이 미이 8 비민수 토계부선 ····································
1) Mann-Whitney II 검정
2) Wilcoxon 부호-서역 건전
3) Kruskal-Wallis H 검정 ··································

I. 데이터 코딩

1. 데이터 코딩 방법

엑셀에서 코딩할 경우는 오름차순 숫자로 ID(예: 1, 2, 3...) 설정하고 하나의 설 문지에 대한 응답은 한줄로 입력해야 합니다.

Ca	1	(°1 -) ±			_	3	코딩샘플 - M	icrosoft Excel		-					3
100	· ·	삽입 페이	지 레이아웃	수식	데이터 경	검토 보기									×
[불여 *	· · · · · · · · · · · · · · · · · · ·	다운고딕 가 가 <u>가</u> ~	- 11	· 가 가 가 가	= = ,		하 일반 펌- ₩-	% ,	· 8 조건부 서식 ·	표 셀	급 = 삽입 - 감* 삭제 -	Σ - <mark>10</mark> 	및 찾기 및 데 · 선택 ·		
클립	보드 영		글꼴	rş.	5	방충	15 II	시 형식	G	스타일	셑	1	편집		
	D25	-	(•	f _x											×
	A	В	С	D	E	F	G	Н	I	J	К	L	М	N	-
1	ID	연령	성별	결혼상태	종교	최종학력	건강상태	스트레스1 :	스트레스2	스트레스3 스	·트레스4 스	느트레스5 -	스트레스6	스트레스7	
2	1	1	1	. 1	4	2	3	3	0	3	3	1	2	1	
3	2	2	2	2	1	2	3	3	2	3	3	3	3	1	=
4	3	4	2	2	4	2	3	3	1	3	3	3	3	1	
5	4	3	1	. 1	1	2	2	4	3	4	4	4	4	0	
6	5	2	1	. 1	3	2	2	0	1	0	0	0	0	0	
7	6	1	2	2	4	2	2	0	0	0	0	0	0	0	
8	7	1	2	2	1	2	3	1	0	0	0	1	0	0	
9	8	3	1	. 1	1	2	3	1	3	1	0	0	0	0	
10	9	3	2	1	4	2	3	4	1	2	1	0	0	0	
11	10	4	1	. 1	3	2	2	2	0	1	1	1	0	0	
12	11	4	2	1	4	2	2	2	1	1	1	2	0	0	
13	12	2	1	1	1	2	2	2	2	2	2	1	2	0	
14	13	1	2	1	1	1	3	1	0	1	0	0	1	1	
15	14	3	2	1	4	1	2	1	2	0	2	2	1	0	
16	15	1	1	. 2	3	1	1	1	1	2	1	1	2	1	
17	16	2	2	2	1	3	2	2	0	2	1	1	1	1	
18	17	4	2												
19	18														-
14 4	► ► Shee	et1 / Sheet2	2/Sheet3	20/	() ()		in in				ш. 			•	ī
준비												100%	Θ) .::

한글이나 워드에서 코딩을 할 경우 또한 설문지 한 장에 대한 응답이 뛰어쓰기 없 이 한줄로 입력해야 합니다. (아래 그림은 위 코딩 데이터를 한글에서 작업한 결과)

2	빈 문서 1 - 한글과컴퓨터 초글		×
j I	파일(E) 편집(E) 보기(U) 입력(D) 모양(J) 도구(K) 표(Q) 창(W) 도움말(H)		×
1 0	li 🗇 🗁 • 🗂 🗁 💷 🖷 🗠 • 🐠 🐎 🗈 🖆 🖉 🚿 Z. 🕺 💷 🔗 👒 🝙 🖽 • 🌆 🐋 🎋 • 🗠 🗅	모 😐	Ŧ
	토바탕글 · '슈대표 · '亦바탕 · ' 카 10 · ' 차 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가		Ŧ
1 -	그리기 - ヽ ㅁ ㅇ ヽ d ᆼ / 옷 ヽ ㅍ ๏) 딱 ヾ 밖 같 - 앞 - ㅋ ㅋ = - ㅋ - ㅋ ㅋ ㅋ ㅋ ㅋ ㅋ ㅋ ㅋ ㅋ ㅋ		Ŧ
	The second s		F
			1
-10			
5			
	1114233033121		
	2221233233331		
	4224233133331		
ca-	3111224344440		
	2113220100000		
m-1	1224220000000		
	1221231000100		
	3111231310000		
<u>~</u>	3214234121000		
4	4113222011100		
10-	4214222111200		*
	2111222222120		*
e			=
	·····································		1

복수응답에 대한 팁을 드리면, 복수응답의 경우 엑셀의 경우 문항이 4개라면 4개 의 칸을 만들어서 응답이 되어 있는 것은 모두 입력해야 되며 한글의 경우 4개의 칸을 모두 채워야 합니다.

ex) 3200←(응답이 3과 2인 경우), 1240, 3324 여기서 0은 무응답으로 칸을 채우는 역할을 합니다.

Ca		*	(~ -) =			
	홈	Å	삽입 페이	지 레이아웃	수식	데이터 경
문 붙여, 클립	월 기 성기 정보드 □	망 기	t은고딕 ᆘ <i>가</i> <u>가</u> ▼	+ 11 	· 가 가 <u>가</u> · 배취 ·	
	D5	i	Ŧ	()	fx	
	А		В	С	D	E
1	복수응답	1	복수응답2	복수응답3	복수응답4	
2		3	2	0	0	
3		1	2	4	0	
4		3	3	2	4	
5						
6						

사전-사후 또는 반복측정에 대한 데이터 코딩의 경우 사전-사후의 응답은 같은 줄에 있어야 합니다.

0		~ @ v) ₹			1	코딩샘플 - Mic	rosoft E	xcel	(ten - 10))	1.000	a inner
	<u>-</u>	삽입	페이지 레이아	웃	수식 더	이터	검토	보기				
		맑은 고	딕 - :	11 -	JÎ JÎ			일반	•	國 조건부 서 國 표 서식 ▼	식 *	음··· 삽 음*· 삭
붙이	부봉기 🥩)	2E - 3	· - <u>-</u>	▼ 내천 ▼			*.0 .0	8	🥑 셀 스타일	+	間서
클럽	입보드 🕫		그문		Tái -		맞춤 😡	표시 형	명식 대	스타일		셸
	B7		• (?	f _x								
	A		В		С		D			E		F
1	ID	사전	_몸무게	사전	신장	-	사후 <mark>_</mark> 몸무게		사후_신	신장		
2		1	65			170		64		170		
3		2	80			175		74		175		
4		3	45			154		55		162		
5		4										
6		5		5								
7	1											
8	-											
9												
10	► N Sh	eet1 /	Sheet2 Shee	t3 ⁄ 🤋]/	_		_	14			1
준비										E		10

2. 엑셀과 한글 데이터 SPSS로 옮기기

1) 엑셀테이터 SPSS로 옮기기

A second	제목업		터집합1] -	IBM SPSS Sta	atistics Dat	a Editor		
	파일(F)	편집(<u>E</u>)	보기(V)	데이터(D)	변환(T)	분석(<u>A</u>)	다이렉트 마케	틩(<u>M</u>)
	, ₩ I	파일(<u>N</u>)			Þ	Te	Ł 🗐	12
	열기(Q) 🕨 🕞 데이터(A)							
	데이	티터베이스	열기(<u>B</u>)		*	🗃 명령문(<u>S</u>)		
	(副 텍스	일 데이터	읽기(<u>D</u>)			금 출력	[결과(<u>O</u>)	
	Cog	gnos CII O I E	남 읽기			급 스크	- 립트(<u>C</u>)	
	Im 닫기(C) Ctrl+F4							
	- 제집	8(<u>S</u>)		Ctrl+S				
11								

위 그림처럼 파일(F) → 열기(O) → 데이터(A)를 선택하여 클릭하면 아래와 같은 창이 열립니다.

🕼 데이터 열	7	×
찾아보기: [] spss_c (1) 코딩샘	▶ 데이터샘플	
파일 이름:	코딩샘플.xlsx	열기(<u>0</u>) 보여너기(P)
비교 규정. 인코딩(E):	보근 파럴 (*.*) Excel (*.xls, *.xlsx, *.xlsm) Lotus (*.w*) Sylk (*.slk) dBase (*.dbf) SAS(*.sas7bdat, *.sd7, *.sd2, *.ssd01, *.ssd04, *.xpt) Stata (*.dta) 테스트(*.bt, *.dat, *.csv)	후 (1) 취소 도움말(H)
	모든 파일 (*.*)	

데이터가 있는 폴더로 이동 후 파일유형에서 "Excel 파일" 또는 "모든 파일"을 선택하면 엑셀 테이터 코딩 파일을 찾을 수 있습니다. 선택해 주시고 "열기(O)"를 클릭하시면



(주의: 코딩된 워크시트 이름을 꼭 확인해야 합니다.)

"Excel 데이터 소스 열기"가 생기고 여기서 확인을 클릭하시면

🕼 *제목업	*제목없음3 [데이터집합2] - IBM SPSS Statistics Data Editor										
파일(F)	편집(E) 보기(V)	데이터(<u>D</u>)	변환(T) 분석(/	<u>)</u> 다이렉트미	ŀ케팅(<u>M</u>) 그2	배프(<u>G</u>) 유틸리	EI(U) 창(W)	도움말(<u>H</u>)			
		🕽 🗠 🦯	- 1	▙=		*	5	3 🛄 1		ABS	
										H.	시: 14 / 14 변수
	ID	연령	성별	결혼상태	종교	최종학력	건강상태	스트레스1	스트레스2	스트레스3	스트레스4
1		1	1	1	4	2	3	3	0	3	-
2		2 2	2	2	1	2	3	3	2	3	
3		3 4	2	2	4	2	3	3	1	3	
4		4 3	1	1	1	2	2	4	3	4	
5		5 <mark>2</mark>	1	1	3	2	2	0	1	0	
6		i 1	2	2	4	2	2	0	0	0	
7		1	2	2	1	2	3	1	0	0	
8		3	1	1	1	2	3	1	3	1	
9		3	2	1	4	2	3	4	1	2	
10	10) 4	1	1	3	2	2	2	0	1	
11	1	4	2	1	4	2	2	2	1	1	
12	12	2 2	1	1	1	2	2	2	2	2	
13	13	1	2	1	1	1	3	1	0	1	
14	14	3	2	1	4	1	2	1	2	0	
15	15	1	1	2	3	1	1	1	1	2	•
16	16	5 2	2	2	1	3	2	2	0	2	
17	17	4	2				1		-	-	
18	18			4			<u>_</u>		-		

엑셀에서 코딩 작업한 내용이 SPSS로 옮겨집니다.

2) 한글 또는 메모장 데이터 SPSS로 옮기기

한글 또는 메모장에서 작성된 코딩 데이터는 엑셀 데이터를 SPSS로 옮기는 방법보 다는 조금 더 복잡하다는 단점이 있지만, 샘플 수가 많다면 데이터를 입력하는데 있 어서 엑셀보다는 한글 또는 메모장이 더욱 시간을 절약할 수 있다는 장점이 있습니 다.

2	반문서 1 - 한글과컴퓨터 훈글	
ĮΠ	일(E) 편집(E) 보기(U) 입력(D) 모양(J) 도구(K) 표(Q) 창(⊻	Ø
1 [🙃 🗁 · 💾 🖨 🖻 🖷 🗠 · 🗠 · 🙈 % 🗈 🛅 🖉	2
1 3	바탕글 🎽 🖧 대표 🎽 🏋 바탕 📑 🕺 카 10 🎽 카	-
1 -	III・ヽロロヽdG 🖉 ∿・国 🐻 🔼 🗽	7
-		
1	1114233033121	
	2221233233331	
	4224233133331	
	3111224344440	
4	2113220100000	
~	1224220000000	
-	1221231000100	
*-	3111231310000	
	3214234121000	
S-1	4113222011100	
	4214222111200	
	2111222222120	
~	1211131010011	
4	3214121202210	
∞-	1123111121121	
4	2221322021111	
<u></u>		
144 1] 비 모 μ 1	
1	1 쪽 1다 · 1 중 16카 · 무다 나눔	I
2. SV		5





제목없	(음2 [데이티	터집합1] -	IBM SPSS Sta	atistics Dat	a Editor		
파일(F)	편집(E)	보기(V)	데이터(D)	변환(<u>T</u>)	분석(<u>A</u>)	다이	
새 ፲	파일(<u>N</u>)			Þ		Ŀ	
열기	열기(0)						
데이	터베이스	열기(<u>B</u>)		•			
() 텍스	변	수					
Cog	Cognos 데이터 읽기						

파일(F) → 텍스트 데이터 읽기(D) 클릭

🚺 데이터 열	71	X
찾아보기:	🐌 데이터샘플 💦 🔽 🔝	
메모장	_코딩샘플.txt	
파일 이름:	메모상_코닝챔플.txt	<u><u><u>g</u>기(0)</u></u>
파일 유형:	텍스트(*.txt, *.dat, *.csv)	붙여넣기(P)
인코딩(E):	로컬 인코딩 🗸	취소
	리포지토리에서 파일 검색(<u>R</u>)	도움말(<u>H</u>)

데이터 열기창에서 → 메모장으로 작성되 코딩샘플을 선택해서 "열기(O)"클릭

🕼 텍스트 가져오기 마법사 6단계	중 1단계 🛛 🔍
628 840 1 81 28 5 630 2400 0 73 40 33 632 10200 0 83 31.08 633 870 0 93 31.17 635 17401 83 41.91	텍스트 가져오기 마법사 사용을 환영합니다! 이 마법사를 사용하여 텍스트 파일에 있는 데이터를 읽고 변수에 관한 정보를 지정할 수 있습니다.
wart wart wart w 1 1 1 1 1 2 1 1 1 1 1 3 3 1	텍스트 파일이 사전 정의된 형식과 일치합니까? ⓒ 예(Y) ⓒ 아니오
텍스트 파일: E:\통계분석\포트클 010.	출리오\데이터샘플\메모장_코딩샘플.txt 20
1 1114233033121 2 2221233233331 3 4224233133331 4 3111224344440	-
5 2113220100000 6 1224220000000 7 1221231000100 8 3111231310000 9 3214234121000	
9 3214234121000 10 4113222011100 11 4214222111200	
< 뒤로(B)	다음(N) > 마침 취소 도움말

텍스트 가져오기 마법사 6단계 시작 → "다음(N)" 클릭

· 텍스트 가져오기 마법사 - 6단계 중 4단계(고정 너비로 배열)						
각 변수가 시작되는 위치를 지정하십시오. 첫 번째 열은 열 0입니다.						
변수 구분선을 삽입하려면 눈금자 또는 데이터 영역에서 원하는 위치를 클릭하십시오. 또는 화살표 키를 사용하여 위치로 이동하거나 열 번호를 입력한 다음 구분 삽입 단추를 누르십시오.						
변수 구분선을 이동하려면 새 위치로 끄십시오.						
변수 구분선을 삭제하려면 위치를 선택하거나 입력하십시오. 그런 다음 삭제 키 또는 구분 삭제 단추를 누르십시오.						
±≡ μ 10 20 30 40 50 60 70						
1 11114233033121 2 223123323331 3 422423133331 4 3111243434440 5 3113240100000 6 122420000000 7 1221231000100 8 3111231310000 9 2114234121000						
열 변호(C): 5 구분 삽입(S) 구분 삭제(D) 현재 변수 너비(V): 1						
< 뒤로(B) 다음(N) > 마첨 취소 도움말						

텍스트 가져오기 마법사 3단계까지 다음을 클릭하면 4단계가 나옵니다.

여기서 마우스로 구분선을 삽입해야 합니다. 코딩샘플은 모두 한자리이지만, 혹시 연령이나 몸무게와 같이 2자리 이상일 때는 두칸 이상으로 구분선을 넣어주어야 합 니다. 구분선을 모두 넣은 후

9411 1 52 2 63	var2 var3 8 840 1 8 2400 0		트 파일의 형식 음에 사용할 수 예(Y) 마니오	을 정의하였습 있도록 이 파일	(니다. ! 형식을 저장하 다른 0	H시겠습니까 비름으로 저장	? }(<u>S</u>)
	2 10200 0 3 870 0 리보기	· · · · · · · · · · · · · · · · · · ·	령문을 붙여넣으 예(Y) 마니오(N) 트 가져오기 미	2시겠습	4분적으로 데이 1면 종료를 누르	E터 캐쉬(ⓒ) ≧십시오.	
V1	V2	V3	V4	V5	V6	V7	
V1 1	V2	V3	V4 4	V5 2	V6 3	V7 3	4
V1 1 2	V2 1 2	V3 1 2	V4 4 1	V5 2 2	V6 3 3	V7 3 3	-
V1 1 2 4	V2 1 2 2	V3 1 2 2	V4 4 1 4	V5 2 2 2	V6 3 3 3	V7 3 3 3	-
V1 1 2 4 3	V2 1 2 2 1	V3 1 2 2 1	V4 4 1 4 1	V5 2 2 2 2 2	V6 3 3 3 2	V7 3 3 3 4	-
V1 1 2 4 3 2	V2 1 2 2 1 1 1	V3 1 2 2 1 1	V4 4 1 4 1 3	V5 2 2 2 2 2 2 2	V6 3 3 3 2 2 2	V7 3 3 3 4 0	4
V1 1 2 4 3 2 1	V2 1 2 2 1 1 2 2	V3 1 2 2 1 1 2	V4 4 1 4 1 3 4	V5 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2	V7 3 3 4 0 0	
V1 2 4 3 2 1 1	V2 1 2 2 1 1 2 2 2 2 2 2	V3 1 2 2 1 1 2 2 2 2	V4 4 1 4 1 3 4 1	V5 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3	V7 3 3 4 0 0 1	
V1 1 2 4 3 2 1 1 3	V2 1 2 2 1 1 2 2 2 2 2 1	V3 1 2 1 1 2 2 2 2 2 1	V4 4 1 4 1 3 4 1 1	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3	V7 3 3 4 0 0 1 1	
V1 1 2 4 3 2 1 1 3 3	V2 1 2 1 1 2 2 1 2 2 1 2 2 1 2	V3 1 2 1 1 2 1 2 2 2 1 1 1	V4 4 1 4 1 3 4 1 1 1 4	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3 3	V7 3 3 4 0 0 1 1 4	
V1 1 2 4 3 2 1 1 3 3 4	V2 1 2 1 1 2 2 2 2 1 2 2 1 2 2 1	V3 1 2 1 1 2 2 1 2 2 1 1 1 1 1	V4 4 1 4 1 3 4 1 1 4 3 3	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3 3 2	V7 3 3 4 0 0 1 1 4 2	
V1 1 2 4 3 2 1 1 3 3 4 4 4	V2 1 2 2 1 1 2 2 2 1 2 2 1 2 2 1 2 2 1 2 2	V3 1 2 2 1 1 2 2 2 1 1 1 1 1 1	V4 4 1 4 1 3 4 1 1 4 3 4 3 4	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3 3 2 2 2	V7 3 3 4 0 0 1 1 4 2 2 2	
V1 1 2 4 3 2 1 1 3 3 4 4 4 2	V2 1 2 2 1 1 2 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 1 2 2 1 1 2 2 1 1 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 2 1 1 2 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2	V3 1 2 2 1 1 2 2 2 1 1 1 1 1 1 1	V4 4 1 4 1 3 4 1 1 4 3 4 3 4 1	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3 3 2 2 2 2 2	V7 3 3 4 0 0 1 1 4 2 2 2 2 2	
V1 1 2 4 3 2 1 1 3 3 4 4 4 2 1	V2 1 2 1 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2	V3 1 2 1 1 2 1 2 2 1 1 1 1 1 1 1 1 1	V4 4 1 4 1 3 4 1 1 4 3 4 3 4 1 1 1	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3 3 2 2 2 2 2 2 3 3	V7 3 3 4 0 0 1 1 1 4 2 2 2 1	
V1 1 2 4 3 2 1 1 3 3 4 4 4 2 1 2	V2 1 2 1 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 2 1 1 2 2 1 1 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2	V3 1 2 2 1 1 2 2 1 1 1 2 1 1 1 1 1 1 1 1	V4 4 1 4 1 3 4 1 1 4 3 4 1 1 4 1 1 1	V5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	V6 3 3 2 2 2 2 3 3 3 3 2 2 2 2 2 2 2 3 3 3 2 2 2 2 2 3 3 3 2 2 3 3 3 2 2 2 2 2 2 3	V7 3 3 4 0 0 1 1 4 2 2 2 2 2 1 1	

4단계 이후 6단계 마지막 단계까지 간 이후 "마침"을 누름

*제목	없음4 [데이티	집합3] -	IBM SPSS St	atistics Data Edit	or		Mag. 28	10.00		1880		• ×
파일(F)	편집(E)	보기(V)	CIIOIEI(D)	변환(<u>T</u>) 분석(A) 다이렉트미	케팅(<u>M</u>) 그라	I프(<u>G</u>) 유틸리티	티(U) 창(W)	도움말(<u>H</u>)			
2				~			*5		2	14 0 (A	
											王人	:13/13변수
		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
1		1		1 1	4	2	3	3	0	3	3	4
2		2	3	2 2	1	2	3	3	2	3	3	
3		4	2	2 2	4	2	3	3	1	3	3	
4		3		1 1	1	2	2	4	3	4	4	
5		2		1 1	3	2	2	0	1	0	0	
6		1		2 2	4	2	2	0	0	0	0	
7		1	1	2 2	1	2	3	1	0	0	0	
8		3		1 1	1	2	3	1	3	1	0	
9		3		2 1	4	2	3	4	1	2		
10		4			3	2	2	2	0	1	ा 	
11		4		1 1	4	2	2	2	2	2	2	
12		2		1 2 1	1		2		2	2	2	
14		3		2 1	4	1	2	1	2	0	2	
15		1		1 2	3	1	1	1	1	2	1	
16		2		2 2	1	3	2	2	0	2	1	
17						6	1.0					
18												

위와 같이 SPSS로 데이터가 옮겨 간 것을 확인할 수 있습니다.

여기서 엑셀은 변수명까지 모두 SPSS로 옮겨지지만 텍스트(메모장)으로 불러온 데이터 는 다시 변수명을 지정해 주어야 합니다. (예: V1→연령)

🚺 *제목입	늘 *제목없음4 [데이터집합3] - IBM SPSS Statistics Data Editor										
파일(F)	편집(E) 보기(V)	데이터(D)	변환(T) 분석	(A) 다이렉트	트마케팅(<u>M</u>) 그래	프(<u>G</u>) 유틸리티	l(U) 창(W)	도움말(<u>H</u>)			
			1		H M					ABG	
	이름	유형	너비	소수점이	설명	값	결측값	엽	맞춤	측도	
1	연령	숫자	1	0		없음	없음	8	📰 오른쪽	臱 명목(N)	N *
2	성별	숫자	1	0		없음	없음	8	🗃 오른쪽	💦 명목(N)	N
3	결혼상태	숫자	1	0		없음	없음	8	🗃 오른쪽	💑 명목(N)	2
4	종교	숫자	1	0		없음	없음	8	🗃 오른쪽	🙈 명목(N)	N
5	V5	숫자	1	0		없음	없음	8	🗃 오른쪽	💑 명목(N)	>
6	V6	숫자	1	0		없음	없음	8	🗃 오른쪽	💑 명목(N)	N
7	V7	숫자	1	0		없음	없음	8	🗃 오른쪽	욿 명목(N)	1
8	V8	숫자	1	0		없음	없음	8	🗃 오른쪽	💑 명목(N)	8
9	V9	숫자	1	0		없음	없음	8	🗃 오른쪽	臱 명목(N)	N
10	V10	숫자	1	0		없음	없음	8	🗃 오른쪽	臱 명목(N)	N .
11	V11	숫자	1	0		없음	없음	8	🗃 오른쪽	💑 명목(N)	N
12	V12	숫자	1	0		없음	없음	8	🗃 오른쪽	臱 명목(N)	N
13	V13	숫자	1	0		없음	없음	8	🗃 오른쪽	🚴 명목(N)	
4.4	1						New York Commence				
데이터 브	[1] [1] [1] [1] [1] [1] [1] [1] [1] [1]										

변수명은 "변수보기(V)"을 클릭하면 넣을 수 있습니다.

또한 변수추가를 통해 ID를 넣을 변수자리를 만들어 ID에 순차적으로 숫자를 넣어주면 엑셀에서 가져 온 데이터와 같은 형태가 됩니다.

📬 *제목없음4 [데이터집합3] - IBM SPSS Statistics Data Editor									2 23			
파일(F)	편집(E) 보기(⊻)	데이터(D) 형	변환(<u>T</u>) 분석	(A) 다이렉트	마케팅(Μ) 그래?	프(<u>G</u>) 유틸리티	l(U) 창(W) .	도움말(번)			
🔁 🖶 🖨 📖 🖛 🛥 🎬 📥 💷 🏴 🛍 🎫 🚟 🐴 🚟 🛶 💊 🧠 🥙												
	이름	유형	너비	소수점이	설등	퀑	값	결측값	열	맞춤	측도	
1	면령	숫자	1	0			없음	없음	8	📰 오른쪽	🚴 명목(N)	N 🖆
2	성별	숫자	1	0			없음	없음	8	遭 오른쪽	🚴 명목(N)	>
3	결혼상태	숫자	1	0	_		없음	없음	8	코 오른쪽	🚴 명목(N)	N
4	종교	숫자	1	0	1	변수값	설명	28		10.000	x)	>
5	V5	숫자	1	0		LH 4 71	4 PH				>	>
6	V6	숫자	1	0		-면수값 ·				_		
7	V7	숫자	1	0		기준값((<u>A</u>):			맞춤팁	₫(S))	
8	V8	숫자	1	0		설명(<u>L</u>)	:)	N
9	V9	숫자	1	0			1 = "=	1성")	N
10	V10	숫자	1	0		2	2 = "D					N
11	V11	숫자	1	0		HE						N .
12	V12	숫자	1	0			1 2 11 2 2					
13	V13	숫자	1	0		X	[A(<u>R</u>))	
4.6	4	NAMES OF TAXABLE PARTY.	and the second									4
데이터 브	데이터 보기(V) 변수 보기(V)											
	IBM SPSS Statistics 프로세서 준비 완료											

성별의 값(노란색)을 클릭하면 "변수값 설명"이라는 창이 뜨고 여기에

기준값 칸 1→설명칸 남성이라고 넣고 "추가"버튼을 누르면 남성이 추가됩니다.

이런식으로 기준값 칸 2→설명칸 여성이라고 넣고 "추가"버튼을 누르고 "확인"버튼을 누르면 변인값을 넣을 수 있습니다.

아래 결혼상태의 경우 1→미혼, 2→기혼 이라고 넣으면 됩니다.

3. 역코딩

설문에 대한 척도를 보면 대게 리커트 척도로 이루어진 설문의 경우 역으로 물어보 는 문항이 존재합니다. 이런 문항은 역코딩을 해줘야 합니다.

이 작업을 실수로 지나쳐 버린다면, 아마도 통계분석을 다시해야 할 경우가 생길수 도 있는 매우 중요한 작업입니다.

예) 스트레스1 문항은 5점 척도로 이루어졌으며, 역문항이었음, 결측값은 없었음.

ta *엑셀_데이터샘플.sav [데이터집합1] - IBM SPSS Statistics Data Editor										
파일(F) 편	[집(E) 보기(V)	데이터(D)	변환(<u>T</u>)	분석(<u>A</u>)	다이렉트 마	케팅(<u>M</u>)	그래	프(<u>G</u>) 유틸리	비티(<u>U</u>) 창(<u>W</u>)	도움말(<u>H</u>)
			📑 변수	┝계산(<u>C</u>) □ 스 ા⊧의 기	·비도(O)			*5		1
4 : 스트레스	1 4		210	까드 대극 및 게도/E)	(210)					
	ID	연령					-	최종학력	건강상태	스트레스1
1	1		2 같은	는 면수도 그	28면경(<u>5</u>)			2	3	3
2	2		🔤 다른	른 변 수 로 코	된 변경(<u>R</u>)			2	3	3
3	3		🛐 자동	통 코딩변경	(<u>A</u>)			2	3	5
4	4		11 1	2월 빈 만들	71(<u>B</u>)			2	2	4
5	5		🎶 최적	법의 빈 만들	₹7Im			2	2	1
6	6		р÷	비하를 위한	·데이터 주비/	P)		2	2	1
7	7			의미 시계성		0	-	2	3	1
8	8		H 문주	이번수 생성	(<u>K</u>)			2	3	1
9	9		🗎 날찌	대및 시간 [])법사(<u>D</u>)			2	3	4
10	10		<u>М</u> И Я	예열변수 생	성(<u>M</u>)			2	2	2
11	. 11		및 3 결측	특값대체(⊻)			2	2	2
12	12		🝘 단각	≥생성기(G	i)			2	2	2
13	13				-	OHLO		1	3	1
14	14		비행 번 전	2 8 M (L)		GIII+G		1	2	1
15	15	1		1	2		3	1	1	1
16	16	2	2	2	2		1	3	2	2

변환→ 같은 변수로 코딩변경 또는 다른 변수로 코딩변경 저희는 같은 변수로 코딩변경을 하겠습니다.



스트레스1 문항을 "숫자변수(V)"로 옮기고 "기존값 및 새로운 값(O)"를 클릭합니다.

-기존값	─────────────────────────────────────
 ⑥ 값(♥): 	◎ 기준값(A): 1
5	◎ 시스템-결측값(Y)
 ○ 시스템-결촉값(S) ○ 시스템 또는 사용자 결촉값(U) ○ 범위(N): 에서(T) ○ 최저값에서 다음 값까지 범위(G): 	기존값> 새로운 값(<u>D</u>): 1> 5 2> 4 3> 3 4> 2
 ○ 다음 값에서 최고값까지 범위(E): ○ 기타 모든 값(Q) 	

기존값에 원래 있던 숫자를 넣고 새로운 값에 새로 넣을 숫자를 넣으면 됩니다. 5점 척도 역코딩이니, 1→5로 2→4로..... 5→1로 변환하면 됩니다. 모두 추가를 했다면 계속을 클릭하면 역코딩이 완료되었습니다.

이제 역코딩까지 끝났다면, 요인을 만들어줘야 합니다.

4. 요인 만들기

1) 변수계산(C)를 활용한 요인 만들기

👍 샘플데이	실 샘플데이터.sav [데이터집합1] - IBM SPSS Statistics Data Editor									
파일(F) 된	편집(E) 보기(V) 데이터()) <mark>변환(T)</mark> 분석(A) 다이렉트 마케팅(M)	그래프(G) 유틸리티(U	J) 창(W) 도움말(H)						
		■ 변수계산(C) ☑ 케이스 내의 값빈도(O) 값이동(C)								
	NO	값 이공(_) 같은 변수로 코딩변경(§)	. 연령	연령더미1						
1	1	🌆 다른 변수로 코딩변경(R)	65	50CH .00						
2	2	🔯 자동 코딩변경(A)	78	70대 1.00						
3	3	▶월 비주얼 빈 만들기(B)	84 80CH	이상 .00						
4	4	■ 최저의 비 마들기(n)	73	70CH 1.00						
5	5	· · · · · · · · · · · · · · · · · · ·	70	70CH 1.00						
6	6	모영화들 위안 네이터 준비(P)	65	50CH .00						

변화(T) → 변수계산(C) 클릭

1월 변수 계산 · · · · · · · · · · · · · · · · · ·	×
대상변수(T): 요인이를 = 유형및 설명(L) * 구인거로 * 값1 * 값2	함수 집단(<u>G)</u> :
	모두 산술 CDF 및 비중심 CDF 변환 현재 날짜/시간 날짜 산술 날짜 작성 나파 후호 함수 및 특수변수(E):
응 SE13 응 SE14 응 SE15 응 SE16 ✔ 자기효능감SE총합 ▼	
확인 불며넣기(P) 재설정(R) 취소 도움말	

대상변수(T)에 만들 "요인이름"을 넣고 숫자표현식(E)에 요인을 계산해주면 됩니다. 여기서는 요인이름 = (값1 + 값2 + 값3 + 값4)/4 값1~값4까지 4개의 문항으로 이루어진 "요인이름"이라는 요인을 만드는 작업입니다. 숫자표현식(E)에는 직접입력도 기존값을 마우스로 옮기셔도 상관없습니다.

* : 곱하기, ** : 자승, / : 나누기, ~= : 같지 않다, ~ : ~는 아니다, & : and, | : or

2) 명령문(Syntax)을 활용한 요인 만들기

SPSS에서 요인을 만들수도 있지만, 문항이 많을 경우 매우 작업시간이 오래 걸릴 수 있으니, 이때는 명령문(Syntax)을 활용해 보도록 합니다.



파일(F) → 새 파일(N) → 명령문(S) 클릭

위 그림과 같이 실행하면 아래 그림처럼 Syntax 편집기를 사용할 수 있습니다. 빨간 동그라미 초록색 화살표는 명령어를 실행시키는 버튼입니다.

요인을 만드는 계산식

↓ compute 요인이름 = 계산식. execute.



마우스로 사용할 명령어를 드레그 한 후 실행버튼을 누르면 실행이 됩니다.



스트레스에 대한 평균의 경우 스트레스_평균 = (스트레스1+스트레스2+스트레스3+스트레스4)/4. 스트레스_총합 = 스트레스1+스트레스2+스트레스3+스트레스4. 마지막에 점(.)은 꼭 넣어줘야 합니다.

5. 연속형 데이터를 범주형 데이터로 바꾸기

나이, 키, 몸무게와 같은 연속형 변수를 일반적인 특성 또는 차이분석 변인으로 사용하려면 범주형 데이터로 전환해야 합니다.

🕼 샘플데	🚰 샘플데이터.sav [데이터집합2] - IBM SPSS Statistics Data Editor								
파일(F)	편집(E) 보기(V) 데(이터(D) 변환(T)	분석(<u>A</u>)	다이렉트 마케팅	₽(<u>M</u>)	그래프(G) 유틸리티(U)) 창(<u>W</u>) 도움말		
		0 0	보고	2서(巴)	۲	AA 👯 🚺			
			기술) 동계량(E)	P.	123 빈도분석(F)			
1:NO	1	1	Ŧ		×.	🔚 기술통계(D)			
	NO	sex	평균	문 비교(<u>M</u>)	•		결혼상태		
			일빈	난선형모형(<u>G</u>)	•	₩ 교차부석(C)			
1	1	여성	일빈	!화 선형 모형(Z)	۴.	표····································	기혼		
2	2	남성	혼힡	날모형(X)	×.		기혼		
3	3	남성	상관	·분석(<u>C</u>)	Þ.	<u>ア</u> -ア エ 並(P)	기혼		
4	4	여성	회7	· 분석(<u>R</u>)	۲	<u>정</u> -Q 도표(Q)	기혼		

분석(A) →빈도분석(F) 을 실시



연령(age)를 선택해서 변수로 옮기고 "통계량(S)"를 누르면

빈도분석: 통계량 창이 나옵니다. 여기서 사분위수, 평균, 중위수, 최빈값, 합계를 클릭하여 빈도분석을 실시합니다.

age			
N	유효	146	
	결측	0	
평균		76.28	
중위수		76.00	
최빈값		74	
합계		11137	
백분위수	25	71.00	1
	50	76.00	
	75	82.00	

25분위	평균	71세
50분위	평균	76세
75분위	평균	82세

단위는 60세, 70세, 80세 이상 3 집단으로 나누는 것이 적당함

age

		빈도	퍼센트	유효 퍼센트	누적퍼센트
유효	65	11	7.5	7.5	7.5
	66	3	2.1	2.1	9.6
	67	2	1.4	1.4	11.0
	68	4	2.7	2.7	13.7
	69	4	2.7	2.7	16.4
	70	9	6.2	6.2	22.6

	샘플데이터.sav	[데이터집합2]	- IBM	SPSS	Statistics	Data	Editor	
--	-----------	----------	-------	------	------------	------	--------	--

파일(F)	편집(E) 보기(V)	데이터(D)	변환(T)	분석(<u>A</u>)	다이렉트 마커	∥틩(<u>M</u>)	그래	표(<u>G</u>)	유틸리티(신	J) 창(W)
1:NO	1	r	11 년 수 12 년 수 12 년 (12 년 ([:] 계산(<u>C</u>) 스 내의 굽 비동(F)	 았빈도(<u>O</u>)			*		- C
	NO		교같은	: 변수로 3	18년경(<u>S</u>)				종교	결혼:
			M 나는	:면수도 그	19면경(<u>R</u>)				2152	2
1		1	[🛐 자동	· 코딩변경	(<u>A</u>)				기족교	
2		2		언비마들	≣7I/B)				기독교	
3		3			≥ / (<u>U</u>)				기독교	
4		4	🔣 최적	1의 빈 만들	≣7I(I)				없음	
5		5	모혈	화를 위한	! 데이터 준비(P)	*		기독교	
6		6	🛃 순위	I변수 생성	(<u>K</u>)				천주교	
7		7	🔒 날찌	F및 시간 [과법사(<u>D</u>)				기독교	
8		8	🗖 시경	역벼스 생	년(M)				불교	
9		9	50.21 ×	211110	n				기독교	
10		10		ταλ μη ΛΠ(<u>Μ</u>	.)				기독교	
11		11	🎯 난수	≃생성기(@	<u>3</u>)				기독교	
12		12	● 변환	! 중지(T)		Ctrl+G			천주교	

이제 연속형을 범주형으로 만들기 위해 변환(T)→다른 변수로 코딩변경(R)을 클릭합니 다.



연령(age)를 오른쪽으로 옮긴 후 출력변수 이름(N)에 "범주형_연령"이라고 새로운 이름 을 넣고 "기존값 및 새로운 값(O)"를 클릭합니다.



69세 이하를 1번으로 (여기 데이터에서 60세 이상만 있었음) 70~79세를 2번, 80세 이상을 3번으로 정해줍니다. 기존값에서 범위를 새로운 기준값 에 1, 2, 3를 넣고 계속을 클릭합니다.

		빈돠	퍼센트	유효 퍼센트	누적퍼센트				
유효	1.00	24	16.4	16.4	16.4				
	2.00	74	50.7	50.7	67.1				
	3.00	48	32.9	32.9	100.0				
	합계	146	100.0	100.0					

분석(A)→새로만든 "범주형_연령"변수를 선택해서 빈도분석을 실시하면,

1번(60세) 16.4%, 2번(70세) 50.7%, 3번(80세 이상) 32.9%로 나뉘어진 것을 알 수 있 습니다. 이처럼 연속형 데이터를 범주형 데이터로 전환할 때는 각 항목의 %가 너무 작지 않게 지정해주는 것이 좋습니다.

6. 이상값 찾기

데이터를 변환하고 요인을 만들었다면, 이제 분석을 실시하면 됩니다. 하지만 코딩이라는 자체가 사람이 직접하는 부분이라서 입력실수가 발생할 수 있습 니다. 예를 들어 성별의 경우 1=남자, 2=여자로 범위는 1~2로 한정되어 있지만, 코 딩 데이터에 3 또는 4가 들어가 있다면, 이는 잘못 입력된 데이터가 됩니다. 이를 방지하기 위하여, 변수에 대한 오름차순 또는 내림차순을 통해 잘못 입력된 데이터 값이나 이상값을 찾을 수 있습니다.

🕼 *샘플(레이터	sav [데이	티티집합	말2] -	IBM SP	SS Statisti	cs Data	Editor				
파일(F)	편집	(E) 보	7I(V)	데이	IEI(D)	변환([])	분석(<u>A)</u> [H	이렉트	과케팅(<u>M</u>)		
						3				भ		
1:sex			2									
		N	10			sex		age		연렬		
					잘리	ᅫ내기(丁)						
1		1		1				복사(<u>C</u>)				
2				2		붙여넣기(P)						
3				3		지두	우기(E)			80CH		
4				4		 ∏ 2 ⊭∠	~ 사이까					
5				5				, 	0			
6				6		오들	통자운 성	:멸(<u>A</u>)	5			
7				7		내릴	H림차순 정렬(<u>D</u>)					
8				8		기술	· 통계량		15			
9		9			att. 마츠비/이			0				
10				10								
11				11		여성	u u u		75			
40				10		1.67	A-L		<u></u>			

sex에 마우스를 놓고 오른쪽 마우스버튼을 클릭하면,

잘라내기(T), 복사(C) 와 같은 창이 뜹니다.

여기서, 오른차순 정력(A) 또는 내림차순 정렬(D)를 통해 잘못 입력된 값이나, 이상값을 찾으면 됩니다.

Ⅱ. 초급 통계 분석

인구학적 특성의 경우 대부분 빈도분석 및 기술통계로 이루어집니다.

1. 인구학적 특성에 따른 빈도분석의 팁

1 생플데이	터.sav [데이터집합1] -	IBM SPSS Statistics	Data Editor						×
파일(F) !	편집(E) 보기(V) 데()IEI(D) 변환(T)	분석(A) 다이렉트 마케팅	!(<u>M</u>)	그래프(<u>G</u>)	유틸리티(민) 창(<u>W</u>) 도움말(H)	
😂 🖩 🖨 🛄 🗠 🛥		보고서(P) 기술통계량(E)							
			Ħ	•	□□ 통계표 장성(C)			표시:118/118	8변수
	NO	sex	평균비교(<u>M</u>)	•	·····································	답변수군(<u>S</u>)	혼상태	학력	5
1	1	며상	일반화 선형 모형(<u>C</u>)	60	CH	기독교	기혼	고졸	-
2	2	남성	혼합 모형(<u>X</u>)	▶ 70	СН	기독교	기혼	전문대졸 이상	5
3	3	남성	상관분석(<u>C</u>)	D	상	기독교	기혼	중졸	5
4	4	4 여성	회귀분석(<u>R</u>)	▶ 70	CH	없음	기혼	중졸	5
5	5	남성	로그선형분석(<u>O</u>)	▶ 70	СН	기독교	기혼	전문대졸 이상	노
6	6	며성	신경망(<u>W</u>)	<u>ه</u> 60	CH	천주교	기혼	고졸	5
7	7	며성	분류분석(Y)	۶ 6 0	СН	기독교	기혼	중졸	2
8	8	남성	차원 감소(D)	60	60대 불교		기혼	무학	노
9	9	며성	···· [신고(일) 청도(A)	70	CH	기독교	사별	중졸	
10	10	머니	-1(<u>0</u>)	70	сH	기도교	기호	고족	4

인구학적 특성의 빈도분석은 분석(A)→표→통계표 작성을 클릭하면



통계표 작성창이 나오고 여기서 인구학적 특성을 마우스로 드레그해서 오른쪽으로 옮긴 후 "요약 통계량(S)"를 클릭함

🙀 요약 통계:				
선택한 변수: (다중 변수)				
통계량(!):	표시(D):			
가중되지 않은 빈도 🛛 🧲	통계량	설명	형식	소수점
행N%	빈도	빈도	nnnn	0
±N%	열 N %	열 N %	nnnn.n%	1
The set of	요약 통계량(<u>C</u>) 표시(D): 토게랴	AD	م ن	
₩N%	응게공	20	0000	<u> </u>
열 N % 표 N % 부표 N % 레이어 N %	모든함목에 적용	(A) E71	F Sector	:₽

통계량(I) 중 "열N%"를 오른쪽으로 옮기고 "선택한 항목에 적용(S)"을 클릭하면

		빈도	열 N %
sex	남성	37	25.3%
	여성	109	74.7%
연령	60 CH	24	16.4%
	70대	74	50.7%
	80대 이상	48	32.9%
종교	기득교	43	29.5%
	천주교	34	23.3%
	불교	33	22.6%
	기타	3	2.1%
	없음	33	22.6%
결혼상태	기흔	81	55.5%
	이혼/별거	4	2.7%
	사별	58	39.7%
	미흔	3	2.1%

한번에 인구학적 특성의 빈도가 나옵니다.

출력된 아웃풋을 가지고 인구학적 특성을 작성하면 됩니다.

2. SPSS 아웃풋을 엑셀로 보내는 방법

SPSS 아웃풋을 바로 사용하면 좋겠지만 엑셀파일로 아웃풋을 바꾸어서 사용하면 더욱 편하게 아웃풋을 사용할 수 있습니다.



내보내기(E)를 클릭합니다.

🔚 내보내기 출력결과	사용자 경기 #		23
┌내보낼 개체			
○ 모두(A) ○ 모두 표시(V) ● 선택(D)			
2010 2010	8.4(O)-		
		이크브 마들기(C)	
		· · · · · · · · · · · · · · · · · · ·	
🕋 표, 텍스트 및 그래픽을 Excel 2007 이상	위크시트이 위치	마지만 역 다음에(4.)	
VML 기반 파일 형식으로 새 워크북이나	피벗표의 레이어	인쇄 레이머 설정에 유효	(EIIOI볼을
기존 워크북에 대보냅니다. 그래픽은 아크뷰에 표하되니다. 사용한 스 아들	각주와 캡션 포함	04I	
귀그북에 도입됩니다. 사용을 두 있는 그래픽 옵션은 없습니다.	모형 보기	인쇄 설정에 유효(각 모형	열에 대한 모
	1		
	옵션 변경(C)		
파일이름(F):			
E:\통계분석\포트폴리오\데이터샘플\인구학적 특성	2.xlsx		찾아보기(B)
1월 파일 저작		×	
찾아보기: 🍶 데이터챔플			
🕋 인구학적 특성 xlsx			
🗃 인구학적 특성2.xlsx		******	
I 코딩샘플.xlsx		-	
		3	四日之八(110)
파일 이름: 인구학적 특성 xlsx		নেক্ষ্ব(S)	
저장 유형: Excel 2007 미상(*.xlsx)		▲ 취소 — —	

"선택"클릭되어는지 확인하고 파일을 저장하기 위해 "찾아보기(B)"를 클릭한 후 파일 을 저장합니다.

Ca	12 5 6	-) -	인구학적 특성 - Microsoft Excel		
	홈 삽입	페이지 레이아웃 수식	데이터 검토 보기		
붙여	월 7 월 7 월 2 월 2 월 2 월 2 월 2 월 2	고딕 • 11 • 7 7 간 • 표 • 3 · 2 • 1 ·	水 = = 급 일반 · · 第 章 書 超 ····· ····· 第 章 書 超 ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ······ ····· ····· ····· ····· ······ ····· ····· ····· ····· ····· ····· ····· ····· ····· ······ ····· ····· ······ ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ······ ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ····· ······ ····· ····· ····· ····· ···	월 조건부 서식 * 음 로 표 서식 * 음 로 설 스타일 * [1] 스타일	[무잡 译 작 第 서 실
	A1	• (* fx			
	A	В	C	DE	E
16	-				
17					
18	[데이터집합	1] E:₩통계분석₩포트폴a	리오♥데이터샘플♥샘플데이터.sav		
19					
20			빈도	열 N %	
21	sex	남성	37	25,3%	
22		며성	109	74,7%	
23	면령	6ULH	24	16,4%	
24			74	50,7%	
25		80대 이상	48	32,9%	
26	쏭교	기독교	43	29,5%	
27		전수교 급 ㄱ	34	23,3%	
28		· 알 씨 - 기도!	33	22,6%	
29		기타	3	2,1%	
30		없음	33	22,6%	
31	글온상태	기온	81	55, 5%	
32		미온/별기	4	2,7%	
33		사업	58	39,7%	
34		비온	3	2,1%	
35					

엑셀로 옮겨진 아웃풋 내용

예제) 대상자의 인구학적 특성 해석

대상자의 인구학적 특성은 '여성'이 109명(74.7%)으로 '남성' 37명(25.3%)보다 많았다. 연 령은 평균 76.28세로 '70대'가 74명(50.7%)으로 가장 많았으며, 다음으로 '80대 이상' 48 명(32.9%), '60대' 24명(16.4%)순이었다. 종교는 '기독교'가 43명(29.5%)으로 가장 많았으 며, 다음으로 '천주교' 34명(23.3%), '불교'와 '무교' 33명(22.6%), '기타' 3명(2.1%)순이었 다. 결혼상태는 '기혼'이 81명(55.5%)으로 가장 많았으며, 다음으로 '사별' 58명(39.7%), '이혼/별거' 4명(2.7%), '미혼' 3명(2.1%)순이었다.

(N=146)

특성	구분	빈도	%
ਮ ਸ਼	남성	37	25.3%
~8 달	여성	109	74.7%
어 귀 (개)	60대	24	16.4%
건녕(세) 76.28±6.80	70대	74	50.7%
10.2010.09	80대 이상	48	32.9%
	기독교	43	29.5%
	천주교	34	23.3%
종교	불교	33	22.6%
	기타	3	2.1%
	없음	33	22.6%
	기혼	81	55.5%
거호사미	이혼/별거	4	2.7%
결근상대	사별	58	39.7%
	미혼	3	2.1%
	전체	146	100.0%

※ 인구학적 특성으로 차이분석(t-test 또는 분산분석)을 실시할 경우 결혼상태에"이혼/별거"와 "미혼"은 기타로 묶어주는 것이 좋습니다.

<표 21> 대상자의 인구학적 특성

3. 교차분석

🕼 샘플데이터	터.sav [데이터집합5] - I	BM SPSS Statistics	Data Edit	or				
파일(F) 편	집(E) 보기(V) 데이	IEI(D) 변환(T)	분석(<u>A</u>)	다이렉트 마케팅	↓(<u>M</u>)	그래프(<u>G</u>)	유틸리티	(<u>U</u>)
		5	보고	1서(<u>P</u>)	ħ.	H		
			기술	·통계량(<u>E</u>)	•	123 빈도분석	ŧ(E)	
			Ŧ		۴.	🔚 기술통계	(D)	
	NO	성별	평균	원비교(<u>M</u>)	•		 탁색(E)	
1	1	чы	일빈	!선형모형(<u>G</u>)	•	교 교차부스		
2	2	년 (11) - 11/2	일빈	!화 선형 모형(<u>Z</u>)	•	문 비용/P)	100,	H
3	3	남성	혼힡	¦모형(<u>X</u>)			 	
4	4	며심	상관	؛분석(<u>C</u>)	×.			H
5	5	남성	회7	분석 <mark>(R</mark>)	×	<u>정</u> 전-전 군 권	±(Q)	Н
6	6	여성	로그	l선형분석(<u>0</u>)	Þ.	6	5	60CH
7	7	여성	신경	망(<u>W</u>)	•	6	5	60CH
8	8	남성	분류	분석(<u>Y</u>)	•	6	5	60CH
9	9	여성	÷1 0		2	7	0	700

분석(A) → 기술통계량(E) → 교차분석(C) 검정을 클릭합니다.



성별에 따른 연령의 차이를 알아보려고 한다면, 행(W)에는 "성별"을 열(C)에는 "연령" 을 넣고 "통계량(S)" 클릭한 후 "카이제곱(H)"을 체크하고 "계속"을 클릭합니다.

🚺 교차분석		x)	🚺 교차분석: 셀 출력	×
✔ NO ▲ ✔ 성별더미 ✔ age ● 연령더미1 ● 연령더미2 ● 양력 ● 양력 ● 양력 ● 양	행(₩):	정확(X) 통계량(S) 열식(E) 붓스트랩(A)	변도 ♥ 관측빈도(Q) ■ 기대빈도(E) ■ 낮은 빈도 숨기기 미만 5 패센트 ♥ 행(R) ● 열(C) ■ 전체(C) - 정수가 아닌 가중값 ● 셀 수 반올림(N) ● 셀 수 절삭(L) ● 조정 없음(M) 14	Z 검정 열 비율 비교 ● 열 비율 비교 ● 값 조정(분페로니법) 전차 ● 비표준화(U) ● 표준화(S) ● 수정된 표준화(A) 이 케이스 가증값 반올림(W) ● 케이스 가증값 철석(H) 취소 도움말
(L)			PIO	.00

셀(E)를 클릭한 후 셀출력창에 퍼센트에서 "행(R)"클릭

성별 * 연령 교차표

				연령		
			60대	70대	80대 이상	전체
성별	남성	빈도	10	21	6	37
		성별 중 %	27.0%	56.8%	16.2%	100.0%
	여성	빈도	14	53	42	109
		성별 중 %	12.8%	48.6%	38.5%	100.0%
전체		빈도	24	74	48	146
		성별 중 %	16.4%	50.7%	32.9%	100.0%

카이제곱 검정

	값	자유도	점근 유의확를 (양측검정)
Pearson 카이제곱	7.925 ^a	2	.019
우도비	8.239	2	.016
선형 대 선형결합	7.841	1	.005
유효 케이스 수	146		

a. 0 셀 (0.0%)은(는) 5보다 작은 기대 빈도를 가지는 셀입니다. 최소 기대빈도는 6.08입니다.

교차분석 아웃풋을 살펴보면, 70대가 50.7%가 가장 많았으며, 다음으로 80대 이상이 32.9%, 60대 16.4%순이었다. 성별에 따른 차이를 살펴보면, 남성은 60대가 여성보다 많았으며, 여성은 80대 이상이 남성보다 많았다. 카이제곱검정결과 X^2 =7.925, p값은 .019 로 유의수준 .05보다 작아 통계적으로 유의미한 차이가 있었다. 라고 해석할 수 있습니다.

인구학적 특성(성별, 연령, 종교 등등)에 따른 차이분석의 경우 t-test 및 분산분석 (ANOVA)을 주로 합니다.

두 분석 모두 집단에 따른 평균차이가 있는지 검정할 때 사용하며, t-test는 두집 단(예: 남자, 여자) 이하, 분산분석(ANOVA)은 세 집단(예: 1학년, 2학년, 3학년)이 상일 때 사용합니다.

또한 다중회귀분석이 필요한 연구에선 종속변수에 영향을 주는 변수(통제변수, 독 립변수)를 찾아 보정하려고 차이분석을 실시하기도 합니다.

4. 독립 t-test



검정변수에는 종속변수(연속된 변수)를 집단변수에는 변인(sex)을 넣고 "집단정의"를 클 릭한다. 집단 남성=1, 여성=2 로 코딩되어 있으니, 1과 2를 넣습니다.

		집	갈통계량									
	sex	N	평균	표준된	년 전차 표	평균의 [준오차		1	_evene 등	분산 검정 결.	과	
자기효능감	남성	37	47.0811	9.45	568	1.55450		🛪 fi	의확률이	0.05보다큼의	2로	
	여성	109	41.5872	7.69	826	.73736						
			Leve	ene의 등	분산 검정	-		평균의	등일성에 대한	: t-검정		
				-								
								유의확률		차이의	차이의 959	6신뢰구간
					유의확를	t	자유도	유의확를 (양쪽)	평균차	차이의 표준오차	차이의 959 하한	6 신뢰구간 상한
자기효능감	등분산0	가정됨		- .410	유의확를 .523	t 3.533	자유도 144	유의확를 (양쪽) .001	평균차 5.49393	차이의 표준오차 1.55507	차이의 959 하한 2.42021	6 신뢰구간 상한 8.56764

독립 t-test 아웃풋이며, 남성의 자기효능감 평균은 47.08점이고 여성은 평균 41.59점 으로 남성이 여성보다 자기효능감 평균이 높았다. 이제 이 평균 차이가 통계적으로 유의 미한지를 살펴봐야 합니다. 독립표본검정을 보면, t=3.533이고 유의확률이 0.001로 유의 수준 0.05보다 작아 통계적으로 유의미한 차이가 나타나고 있었다.

여기서 t값의 경우 클수록 차이가 많이 납니다. 물론 마이너스(-)값도 절대값으로 클수 록 변인에 따른 평균 차이는 많이 납니다. 그 기준은 대략 t=1.96으로 이 값보다 크면 변인에 따른 평균 차이가 있다고 봐도 됩니다.

변인	구 분	M(평균) ± SD(표준편차)	t값	p(유의확률)
 서 벼	남성	47.08 ± 9.46	2 5 2 2	0.001
~8 钽	여성	41.59 ± 7.70	3.033	0.001

위 아웃풋을 표로 표현하면 위와 같이 됩니다.

5. 대응 t-test

두 집단에 따른 차이가 있다면, 같은 집단 안에 두 측정시점에 따른 평균 차이도 있습니다. 예를들어 한 교실에서 1학기 때의 키와 몸무게, 2학기 때의 키와 몸무게 의 차이와 같이 한 집단에서 시간 차를 두고 평균을 비교하는 부분입니다. 보통 어 떠한 프로그램을 적용하기 전-후 비교로 많이 사용합니다.



사전-사후 비교할 변수를 짝을 맞추어 넣고 확인을 누르면

		년 평	N	표준편차	평균의 표준오차
대응 1	체중_사전	67.550	30	6.7186	1.2266
	체중_사후	65.177	30	7.1128	1.2986
대응 2	제지방량_사전	43.233	30	3.3585	.6132
	제지방량_사후	42.60	30	3.434	.627
대응 3	근육량_사전	40.297	30	3.0824	.5628
	근육랑_사후	39.723	30	3.1726	.5792

대응표본 통계량

사전 체중은 67.55kg에서 프로그램 적용 후 65.18kg로 약 2.37kg 감소하였으며, 대응 표분 검정결과 t값이 5.690로 나타났다. 유의확률이 0.000으로 통게적으로 유의미한 차 이가 나타났다는 것을 알 수 있습니다.

체지방량과 근육량도 위와 같이 해석을 할 수 있습니다.

대응차 차이의 95% 신뢰구간 평균의 유의확를 평균 표준편차 표준오차 하한 상한 자유도 (앙쪽) 대응 1 체중_사전 - 체중_사후 2.3733 2.2844 .4171 1.5203 3.2263 5.690 .000 29 대응 2 제지방량_사전 -.6333 1.2666 .2313 .1604 1.1063 2.739 29 .010 제지방량_사후 근육량_사전 -대응 3 .5733 1.2006 .2192 .1250 1.0216 2.616 29 .014 _ 근육량_사후

대응표본 검정

대응표본검정표에 나타난 평균은 사전-사후 평균차이입니다. 만약 사후에 변수값이 증가하였다면, 마이너스로 나올 것입니다.

변인	구 분	M(평균) ± SD(표준편차)	t값	p(유의확률)	
케즈	사전	67.55 ± 6.72	E 600	0.000	
শান্ত	사후	65.18 ± 7.11	5.090	0.000	
	사전	43.23 ± 3.36	0.720	0.010	
세사방장	사후	42.60 ± 3.43	2.739	0.010	
그으라	사전	40.30 ± 3.08	9.616	0.014	
こずび	사후	39.72 ± 3.17	2.010	0.014	

위 아웃풋을 표로 표현하면 위와 같이 됩니다.

6. 분산분석(ANOVA)

1 생물	들데이터	.sav [데이터집합	6] - IBM SPS	S Statistics	Data Edit	tor						
파일(5) 편집	집(E) 보기(V)	데이터(D)	변환(丁)	분석(<u>A</u>)	다이렉트 마케팅	∃(<u>M</u>)	그래프	.(<u>G</u>) 유	틸리티(<u>U</u>)	창(<u>W</u>) 도움말
				7	보고 기술	1서(P) 5통계량(E)	•	H	*5			
					Ŧ		۲					
		NO		sex	평균	7 비교(M)	×	<u> </u>	단별 평균	분석(M)		상태
					일빈	·선형모형(<u>G</u>)	۲	만 일.	표본工검	 정(S)		
	1		1	여성	일빈	!화 선형 모형(Z)	٢	동	립표본T	_ 검정(T)		기혼
1	2		2	남성	혼힡	:모형(X)	۲		- T 복표응	검정(P)		기혼
	3		3	남성	상관	<u> ·</u> 분석(<u>C</u>)	•		원배치 분	사부선(이)		기혼
4	1		4	여성	회7	· 분석(<u>R</u>)	•		은배지 전			기혼
. !	5		5	남성	로그	1선형분석(<u>O</u>)		70CH		기독교		기혼
(6		6	여성	신경	명망(W)		60CH		천주교		기혼
	7		7	며실	분류	 루분석(Y)		60CH		기독교		기혼
1	3		8	남성	차운	· _ · · · · · ·	- 10 - 10	60CH		불교		기혼
(9		9	여성	처리		0	70CH		기독교		사별
1	0		10	여성	45		1	70CH		기독교		기혼
1	1		11	며심	비노	2수 김성(<u>N</u>)	1	70CH		기독교		기혼

분석(A) → 평균비교(M) → 일원배치 분산분석(O)를 클릭



독립변수를 요인에 넣고 종속변수를 종속변수에 넣고 사후분석(H)를 클릭

<u>L</u> SD	📃 <u>S</u> -N-K	Maller-Duncan
Bonferroni	🔽 <u>T</u> ukey 방법	제1종(제2종 오류 비율(/): 100
Sidak	🔲 Tu <u>k</u> ey의 b	Dunn <u>e</u> tt
Scheffe	🔽 Duncan	제어 범주(Y): 마지막(T) 👻
<u> R</u> -E-G-W의 F	📄 <u>H</u> ochberg의 GT2	[검정
🖪 R-E-G-W의 <u>Q</u>	🗾 <u>G</u> abriel	⑧ 양쪽검정(2) ◎ < 통제(○) ◎ > 통제(N)
등분산을 가정하지 않 Ta <u>m</u> hane의 T2	방음 ☑ Dunnett의 T <u>3</u> ☑] G <u>a</u> mes-Howell 🔲 D <u>u</u> nnett의 C

다양한 사후분석이 나타납니다.

여기서 등분산이 가정되었을 때

Scheffe, Tukey 또는 Duncan을 많이 사용합니다.

이들의 차이는 Duncan와 Tukey의 경우 변인의 수가 비슷할 때 사용합니다. ex) 1학년= 35명, 2학년 =36명, 3학년 34명

Scheffe는 경우 변인의 수가 일정하지 않을 때 사용합니다. ex) 1학년= 45명, 2학년 =26명, 3학년 64명

대부분 이 경우라서 Scheffe 사후검증을 실시하게 됩니다.

등분산이 가정되지 않았을 때

Dunnett T3 또는 Games-Howell를 사용합니다.

"계속"를 클릭하고 "옵션(O)"을 클릭하면

	원배치 분산분석 NO sex age 종교 결혼상태 학력 동거유형1 노인부부만산다 고호한다(40)사단		:변수(E): 자기효능감 (E): 연령 정(R) 취소	IIIII(C) 사후분석(H) 옵션(Q) 봇스트랩(B)	 알원배치 분산분석: ※ 계량 ⑦ 기술통계(D) □ 모수 및 변량효과(F) ※ 분산 동질성 검정(H) □ Brown-Forsythe(B) ♡ Welch(W) □ 평균 도표(M) · 결측값 · ● 분석별 결측값 제외(A)
남성	68	60CH	천주교	사별	◎ 목록별 결측값 제외(L)
며성	68	60CH	천주교	기혼	계속 취소 도움말
남성	71	70 CH	기독교	기혼	

통계량을 체크하는 창이 나옵니다.

여기서는 기술통계(D), 분산 동질성 검정:Levene 통계(H), Welch(W)를 클릭합니다. Welch는 등분산이 가정되지 않았을 때 분산분석표 대신 사용하는 값입니다.

기술통계

자기효능감

					평균에 대한 95% 신뢰구간			
	N	평균	표준편차	표준오차	하한값	상한값	최소값	최대값
60대	24	48.2500	6.96107	1.42092	45.3106	51.1894	34.00	58.00
70대	74	44.2027	7.91215	.91977	42.3696	46.0358	26.00	61.00
80대이상	48	38.4583	8.00255	1.15507	36.1346	40.7820	22.00	59.00
합계	146	42.9795	8.49054	.70268	41.5906	44.3683	22.00	61.00

분산의 등질성 검정

자기효능감

Levene 통계량	df1	df2	유의확률
.413	2	143	.663

일원배치 분산분석

ルコ る 는 관

시기포이	•				
	제곱합	df	평균 제곱	F	유의확률
집단-간	1758.562	2	879.281	14.462	.000
집단-내	8694.376	143	60.800		
합계	10452.938	145			

평균의 동질성 검정

자기효능감

	통계량"	df1	df2	Sig.			
Welch	15.175	2	64.552	.000			
a. 자등으로 F 분배합니다.							

우선 Levene값을 살펴보면, 유의확률이 0.663으로 유의수준 0.05보다 커 등분산을 가 정할 수 있으며, 이 경우 분산분석의 F값을 가지고 해석하면 됩니다.

자기효능감의 경우 60대 평균은 48.25점, 70대 평균은 44.20점, 80대 이상 평균은 38.46점으로 나타났으며, F값이 14.462이고 유의확률이 0.000으로 통계적으로 유의미한 차이가 나타났다. 라고 해석하시면 됩니다.

사후검정

						95% 신뢰구간	
	(I) 연령	(J) 연령	평균차(I-J)	표준오차	유의확를	하한값	상한값
Tukey HSD	60 CH	70대	4.04730	1.83165	.073	2907	8.3853
		80대 이상	9.79167	1.94936	.000	5.1749	14.4085
	70대	60대	-4.04730	1.83165	.073	-8.3853	.2907
		80대 이상	5.74437	1.44509	.000	2.3219	9.1669
	80대 이상	60대	-9.79167	1.94936	.000	-14.4085	-5.1749
		70대	-5.74437	1.44509	.000	-9.1669	-2.3219
Scheffe	60대	70대	4.04730	1.83165	.091	4835	8.5781
		80대 이상	9.79167	1.94936	.000	4.9697	14.6136
	70대	60대	-4.04730	1.83165	.091	-8.5781	.4835
		80대 이상	5.74437	1.44509	.001	2.1698	9.3190
	80대 이상	60대	-9.79167	1.94936	.000	-14.6136	-4.9697
		70대	-5.74437	1.44509	.001	-9.3190	-2.1698
Dunnett T3	60대	70대	4.04730	1.69263	.062	1479	8.2425
		80대 이상	9.79167	1.83118	.000	5.2804	14.3029
	70대	60대	-4.04730	1.69263	.062	-8.2425	.1479
		80대 이상	5.74437	1.47654	.001	2.1608	9.3280
	80대 이상	60대	-9.79167	1.83118	.000	-14.3029	-5.2804
		70대	-5.74437	1.47654	.001	-9.3280	-2.1608
Games-Howell	60대	70대	4.04730	1.69263	.054	0585	8.1531
		80대 이상	9.79167	1.83118	.000	5.3745	14.2089
	70대	60대	-4.04730	1.69263	.054	-8.1531	.0585
		80대 이상	5.74437	1.47654	.001	2.2314	9.2574
	80대 이상	60 CH	-9.79167	1.83118	.000	-14.2089	-5.3745
		70대	-5.74437	1.47654	.001	-9.2574	-2.2314

다중 비교

종속 변수: 자기효능감

*. 평균차는 0.05 수준에서 유의합니다.

사후검증을 살펴보면, 여기서는 등분산이 가정되었으며, 변인의 수가 다르기 때문에 Scheffe를 사용하면 됩니다. 여기서 평균차(I-J)칸의 숫자뒤에 별이 붙은 것이 유의미한 차이가 있다는 표시입니다.

하지만 등분산이 가정되지 않았을 경우 Levene 유의확률이 0.05보다 작을 경우 Dunnett T3나 Games-Howell를 사용해야 합니다.

이를 정리한 표가 "동일 집단군"표입니다.

동일 집단군

			유의수준 = 0.05에 대한 부집단		
	연령	Ν	1	2	3
Tukey HSD ^{a,b}	80대 이상	48	38.4583		
	70대	74		44.2027	
	60대	24		48.2500	
	유의확를		1.000	.058	
Duncan ^{a,b}	80대 이상	48	38.4583		
	70대	74		44.2027	
	60대	24			48.2500
	유의확를		1.000	1.000	1.000
Scheffe ^{a,b}	80대 이상	48	38.4583		
	70대	74		44.2027	
	60 CH	24		48.2500	
	유의확를		1.000	.074	

자기효능감

동일 집단군에 있는 집단에 대한 평균이 표시됩니다.

a. 조화평균 표본 크기 39.467을(를) 사용합니다.

b. 집단 크기가 동일하지 않습니다. 집단 크기의 조화평균이 사용됩니다. ㅣ유형 오차 수준은 보장되지 않습니다.

이를 해석하면 60대와 70대 집단의 자기효능감이 80대 이상 집단보다 높았다는 것을 보여준다고 해석합니다.

Scheffe사후검증의 경우 분산분석 후 p값이 0.05보다 작더라도 그룹의 차이가 나타나지 않을 경우가 있습니다. 이는 Duncan보다 사후검증이 더 민감하기 때문입니다. 이럴 경우 해석은 분산분석(ANOVA)의 경우 차이는 나타났지만, Scheffe 사후검증에서는 차이가 나타나지 않았다고 해석해야 합니다.

사후검정 민감도 Scheffe > Tukey > Duncan 순입니다.

Duncan사후검증이 그룹의 차이가 더 잘 나타난다고 보면 됩니다.
예제) t-test 및 분산분석 해석

t-test와 분산분석(ANOVA)를 통해 자기효능감에 대한 평균 차이를 살펴본 결과를 표로 만들고 해석을 하면 다음과 같습니다.

변인	구 분	$M \pm SD$	F/t	р	Scheffe
성별	남성	47.08 ± 9.46	0 500	0.001	
	여성	41.59 ± 7.70	3.533	0.001	
연령	60대(a)	48.25 ± 6.96			
	70대(b)	44.20 ± 7.91	14.462	0.000	c <ba< td=""></ba<>
	80대 이상(c)	38.46 ± 8.00	5.46 ± 8.00		

인구학적 특성에 따른 자기효능감 차이를 살펴보면, 성별(t=3.533, p<.01), 연령 (F=14.462, p<.001)에서 통계적으로 유의미한 차이가 나타났다. 또한 Scheffe's test를 통해 각 연구대상 변인들 사이의 인식 차이를 파악한 결과 살펴보면 다음과 같다. 성별은 남성이 여성보다 자기효능감이 높았으며, 연령은 60대와 70대 집단이 80대 이상 집단보다 자기효능감이 높은 것으로 나타났다.

7. 요인분석

요인분석은 변수들 간의 상관관계를 통해 관측된 변수들에 영향을 미치고 있는 공 통인자를 찾아내는 분석방법입니다. 보통 척도의 개발과정에서 측정도구의 타당성을 파악하는데 사용됩니다. 변수들을 묶어 요인들을 만드는 것이 목적입니다.

표본수는 100개 이상으로 200개는 적당하고 300개정도면 매우 좋습니다.

요인분석은 독립변수와 종속변수를 한꺼번에 투입하여 요인분석을 하는 것이 좋습 니다.

🐪 요인분	t 요인분석.sav [데이터집합4] - IBM SPSS Statistics Data Editor							
파일(F)	편집(E) 보기(V)	데이터(D) 변환(T)	분석(A) 다이렉트 마케팅()	<u>ת (N</u>	래프(<u>G</u>) 유틸리	리티(U) 창(W)		
			보고서(P) ▶ 기술통계량(E) ▶	H	*5			
l.	ID	만족도1 만족	- 프 , 평균비교(M) ▶	-4	만쪽도5	유명한차		
1		1	일반선형모형(G) ▶		1 1	1		
2	2	2 1	일바하셔형 모형(7)		1 1	1		
3		3 1	호하 모형 (V)	3	3 1	1		
4	4	l 1		2	2 1	1		
5		5 1	상관문적(℃) ▶	2	2 1	1		
6		5 1	회귀분석(<u>R</u>) ▶		1 1	1		
7	7	7 1	로그선형분석(⊙) ▶		1 1	1		
8	8	3 1	신경망(₩) ▶		1 1	1		
9	9) 1	분류분석(Y) ▶	Ę	5 1	1		
10	10) 1	차원 감소(◘) ▶	8	요인분석(F)	1		
11	11	1	척도(A) 🕨	1.15	· · · · · · · · · · · · · · · · · · ·	C) 1		
12	12	2 1	비모수 검정(N) ▶		의 해양 전체 전 대() 최저희 최도배(1		
13	13	3 1	예측(T) ▶		J 최억와 억도법(1		
14	14	1	생존확률(S) ▶		1 2	1		

분석(A) → 차원감소(D) → 요인분석(F) 클릭합니다.



요인분석할 변수들을 선택해서 변수(V)로 옮깁니다.

요인분석 🔀	요인분석: 기술통계
변수(V): ● 만족도1 ● 만족도2 ● 만족도3 ● 만족도4 ● 만족도5 ● 유명한차 ● 외제차 · · · · · · · · · · · · · · · · · · ·	통계량 ♥ 일변량 기술통계(U) ♥ 초기해법(!) 산관행렬 ♥ 계수(C) ● 역 모형(N) ● 유의수준(S) ● 재연된 상관행렬(R) ● 행렬식(D) ● 역-이미지(A) ♥ KMO와 Bartlett의 구형성 검정(K) 계속 취소 도움말
확인 붙여넣기(P) 재설정(R) 취소 도움말	
	0 4

변수를 옮긴 후 기술통계(D)를 클릭한 후 일변량 기술통계(U), 계수(C), KMO와 Bartlett의 구형성 검정(K)를 선택한 후 "계속"을 클릭합니다.



요인추출(E)를 클릭한 후 추출에서 고유값 기준 "1"를 선택합니다. 기본값입니다. "계 속" 클릭 후



요인행렬의 열을 최대한 단순화 시키는 작업 중 가장 많이 사용합니다.

요인분석	3 요인분석: 옵션
변수(V): · ID · 한쪽도1 · 만족도2 · 만족도3 · 만족도4 · 만족도4 · 만족도5 · 유명한차 · 유명한차 · · · · · · · · · · · · · · · · · · ·	기술통계(D) 요인추출(E) 요인회전(T) 요인회전(T) 요인점수(S) 옵션(Q) 옵션(Q) 제수출력형식 ····································

옵션(O)를 클릭 후 요인분석 옵션에서 "크기순 정령(S)"를 클릭 한 후 "계속"을 클릭하 면 아웃풋이 나옵니다.

KMO와 Bartlett의 검정

표준형성 적절성의 Kais	.758	
Bartlett의 구형성 검정	근사 카이제곱	561.386
	자유도	45
	유의확률	.000

Kaiser-Meyer-Olkin의 약자 KMO 통계량은 표본 적합도 평가로 .50 이상이어야 하고 보통 .07이상이면 요인분석에 적합한 표본으로 생각합니다.

• KMO
※ .90 이상: 상당히 좋은 편
※ .70~.79: 좋은 편
※ .50 미만: 받아들일 수 없음

※ .80~.89: 꽤 좋은 편※ .50~.69: 평범한 편

Bartlett의 구형성 검정값은 .05보다 작아야 변수들이 상호독립적으며 변수간의 상관이 없다고 판단합니다.

	초기 고유값			추출	추출 제곱합 적재값			회전 제곱합 적재값		
성분	합계	% 분산	% 누적	합계	% 분산	% 누적	합계	% 분산	% 누적	
1	3.612	36.118	36.118	3.612	36.118	36.118	2.368	23.680	23.680	
2	1.415	14.155	50.273	1.415	14.155	50.273	1.952	19.520	43.200	
3	1.130	11.296	61.568	1.130	11.296	61.568	1.837	18.368	61.568	
4	.818	8.175	69.744							
5	.752	7.520	77.264							
6	.628	6.281	83.545							
7	.564	5.637	89.182							
8	.394	3.945	93.126							
9	.385	3.849	96.976							
10	.302	3.024	100.000							

설명된 총분산

추출 방법: 주성분 분석.

설명된 총분산을 살펴보면, 요인은 3개로 나누어졌으며, 모든 요인 고유치가 1이상인 요 인들만 추출되었습니다. 총 누적분산은 61.568로 전체 분산의 61.568%를 설명하고 있습 니다.



a. 4 반복계산에서 요인회전이 수렴되었습니다.

회전된 성분행렬을 살펴보면, 1요인은 만족도1~만족도5로 이루어졌으며, 2요인은 디자 인과 승차감, 3요인은 유명한차, 외제차, 고급차로 이루어졌습니다. 이제 묶어진 요인들에 대한 요인이름을 붙어야 합니다. 1요인은 "만족도"로, 2요인은 "자동차이미지", 3요인은 "자동차 브랜드"로 명명하면 될 듯합니다.

요인분석에서 가장 핵심은 묶이지 않은 변수의 제거에 있습니다. 이는 묶어야 할 요인이 다른 요인에 있거나, 위에 보이는 요인적재량값이 0.3 미만이거나 다른 요인에 걸쳐 비슷 한 요인적재량값을 가지고 있을 때 제거하면 좋습니다.

예) 2요인에 요인적재량이 0.542이고 3요인에 요인적재량이 0.528일 경우 두 요인에 걸쳐 있기 때문에 변수를 제거하고 다시 요인분석을 돌리면 됩니다.



요인추출 방법 중 고정된 요인 수(X)를 선택할 때는 요인분석을 실시하였지만, 제대로 요인분석이 되지 않는 경우 또는 기존 척도보다 오히려 요인의 수가 적게 나온 경우에 선택하여 사용하면 됩니다.

8. 신뢰도 분석

분석을 하기전에 문항을 묶어서 만든 요인들이 신뢰도가 있는지 알아보기 위해, 즉 일관성이 있는 질문인지를 알기위해 신뢰도 분석을 실시합니다.

크론바 알파(Cronbach α)값이 신뢰도 값입니다.

🕼 샘플더	이터.sav [데이터집합5] - IBM SPS	S Statistics	Data Edi	tor					
파일(F)	편집(E) 보기(V)	데이터(D)	변환(<u>T</u>)	분석(<u>A</u>)	다이렉트 마케팅	∃(<u>M</u>)	그래프(<u>G</u>)	유틸리티(U)	창(<u>W</u>) 도울	
			2	보고 기술	1서(P) 호통계량(E)	•	H		- 3 -	
	NO	:	성별	표 평균	원비교(<u>M</u>)	r F	age	연령	연령	
1		1	며심	일면	[선형보형(<u>G</u>)		6	5 600	LH I	
2		2	남성	일빈	!화 선형 모형(Z)		7	8 700	1H	
3		3	남성	혼힡	よ모형(X)	•	8	4 80CH 014	상	
4		4	며심	상관	·분석(<u>C</u>)	•	7	3 700	-H	
5		5	남성	회7	분석 <mark>(R</mark>)		7	0 700	<u>TH</u>	
6		6	며심	로_	1선형분석(<u>O</u>)		6	5 600	1H	
7		7	며심	신경	명망(<u>W)</u>	•	6	5 600	-H	
8		8	남성	분루	루분석(Y)		6	5 600	-H	
9		9	여성	차원	- 감소(D)		7	0 700	IH I	
10		10	며성	철덕	(A)					
11		11	며성	нс			전되도 :			
12		12	남성	013			 [3] 다차원 확장(PREFSCAL)(U) [3] 다차원척도법(PROXSCAL)(P) 			
13		13	며심	ખા≅		P				
14		14	남성	생견	-퐉 <u>듈(S</u>)	M	🎆 다차원칙	덕도법(ALSCAL)(<u>M</u>)	
				L L S	등응답(U)	- No L	Parate and a second		and a second	

분석(A) → 척도(A) → 신뢰도분석(R)을 클릭합니다.

🕼 신뢰도	분석	**			100.000	23	🚰 신뢰도 분석: 통계량	×
✔ NO ♣ 성별 ♣ 409 ♣ 연령 ♣ 전 명 ▲도 월 문	년 [GD] [GD] [GD] [GD] [[GD] [[] [[] [] [] [] [] [] [] [] [] [] [T T T T T T	항목():	1 2 3 4 5 5 5 7 3 3 ·		E.	다음에 대한 기술통계량	한목내 상관관계(R) 공분산(E) 분산분석표 ④ 지정않음(N) ● F-검정(F) ● Friedman 카이제곱(Q) ④ Cocḥran 카이제곱(H) Tuỵey의 가법성 검정(K)
62	SE5	숫자	11	0		없음	모형(<u>M</u>): 이차원 혼합	▼ 유형(Y)]일치 ▼
63	SE6	숫자	11	0		없음	신뢰구간(♥): 95 %5	의케이스추출 검정값(A): 0
64	SE7	숫자	11	0		없음	 [계속]	취소 도움말
FILOLEL HT								

항목(I)에 요인을 이루는 문항을 모두 넣고 "통계량(S)"을 클릭합니다. 여기서 "항목제거시 척도(A)"를 클릭하고 "계속"→"확인"을 클릭합니다.

신뢰도 통계량

Cronbach의 알파	항목 수
.872	16

				항목이 삭제된 경우
	항목이 삭제된 경우 철도 평균	항목이 삭제된 경우 철도 분산	수정된 항목- 전체 상과관계	Cronbach 악파
SE1	40.49	63 341	546	
SE2	40.50	63 659	533	863
SE3	40.65	61 973	629	859
SE4	40.46	66 1 5 3	352	871
SE5	40.46	64.527	.552	865
SE6	40.00	63.697	.512	.000 1.08
957	40.10	66.211	.520	.004
950	40.25	64.024	.570	.070
3E0 0E0	40.32	64.024	.501	.008.
SE9	40.31	63.470	.511	.864
SE10	40.29	64.648	.445	.867
SE11	40.64	64.134	.469	.866
SE12	40.62	63.023	.525	.864
SE13	40.16	62.630	.595	.861
SE14	39.87	65.052	.462	.867
SE15	40.01	63.538	.538	.863
SE16	40.12	62.228	.614	.860

항목 총계 통계량

본 요인의 크론바 알파(Cronbach alphe = α)값은 .872로 기준으로 보는 .06~.70보다 높아 신뢰도가 있다고 해석됩니다.

만약 크론바 알파값이 0.6~0.7 미만으로 나타났다면, 항목이 삭제된 경우의 크론바 알 파값을 살펴봅니다. 항목이 삭제된 크론바 알파값은 항목이 제외되었을 때의 크론바 알파 값으로 여기서는 SE1이 제거되었을 때 크론바 알파값은 .863이네요.

물론 여기서는 전체적으로 신뢰도가 높기 때문에 제거하지 않지만, 혹 신뢰도가 낮다고 하면 항목을 제거했을 때 높아지는 문항을 찾아 제거하고 요인을 다시 묶어야 합니다.

9. 상관분석

상관분석을 실시하는 이유는 측정변수들의 관계의 방향성(예: 음(-), 양(+))과 관 련성을 알기 위해서입니다.

1월 샘플데이터.sav [데이터집합6] - IBM SPSS Statistics Data Editor								
파일(F) 편	!집(E) 보기(V) 데이	비터(D) 변환(T)	분석(A) 다이렉트 마케팅	1 <u>(M)</u>	그래	프(G) 유틸리티(U	U) 창(<u>W</u>)	
		5 3	보고서(P) 기술통계량(E) 표	•	H			
	NO	sex	평균 비교(M) 일반선형모형(<u>G</u>)	•		종교	결혼상E	
1	1	여성	일반화 선형 모형(Z)		SOCH	기독교		
2	2	남성	혼합 모형(X)	•	70CH	기독교		
3	3	남성	상관분석(<u>C</u>)	•	0 🚮	변량 상관계수(<u>B</u>)		
4	4	여성	회귀분석(<u>R</u>)	•	E			
5	5	남성	로그선형분석(<u>O</u>)	•				
6	6	여성	신경망(<u>W</u>)	۰Ļ	0 -			
7	7	며실			SOCH	기독교		
8	8	남성	고 ··· ⊇ ··· (D)		SOCH	불교		
9	9	чn			7008	미론IC		

분석(A) → 상관분석(C) → 이변량 상관계수(B)를 클릭합니다.



상관관계를 보려는 변수를 오른쪽칸으로 이동 후 "Pearson"과 "유의한 상관계수 별표 시"에 체크가 되어있는지 확인 후 "확인"버튼을 클릭합니다.

		자기효능감	영양	스트레스관리	대인관계
자기효능감	Pearson 상관계수	1	.354	.230**	.382
	유의확를 (양쪽)		.000	.005	.000
	Ν	146	146	146	146
영양	Pearson 상관계수	.354	1	.299**	.353
	유의확를 (양쪽)	.000		.000	.000
	Ν	146	146	146	146
스트레스관리	Pearson 상관계수	.230	.299**	1	.240**
	유의확를 (양쪽)	.005	.000		.003
	Ν	146	146	146	146
대인관계	Pearson 상관계수	.382**	.353	.240**	1
	유의확를 (양쪽)	.000	.000	.003	
	Ν	146	146	146	146

상관계수

**. 상관계수는 0.01 수준(양쪽)에서 유의합니다.

상관관계분석 아웃풋으로

해석하면, 자기효능감은 영양(r=.354), 스트레스관리(r=.230), 대인관계(r=.382)와 양의 상관을 가지고 있다라고 해석합니다.

상관분석은 각각의 변인과 변인에 대한 관계를 상관계수(r)로 표현합니다. 양(+) 이면 정적상관이며, 음(-)이면 부적상관입니다. 각각의 상관계수(r)은 다른 상관계수 (r)에 영향을 주지 않는 독립된 값입니다.

상관계수(r)값은 0.7 이상이면 강한 상관, 0.3~0.7은 중등 상관, 0.1~0.3은 약한 상관이라고 해석합니다.

요인	자기효능감	영양	스트레스관리	대인관계
자기효능감	1			
영양	.354**	1		
스트레스관리	.230**	.299**	1	
대인관계	.382**	.353**	.240**	1

예제) 상관분석 해석

** p<.01

대상자의 자기효능감, 건강증진행위 간 관계를 검증한 결과는 다음과 같다. 자아효능감은 건강증진행위 하위 요인 중 대인관계(r=.382)과 가장 큰 양(+)의 상관을 가지고 있었으며, 다음으로 영양(r=.354), 스트레스(r=.230)순으로 나타났다. 10. 회귀분석

"OO이 OO에 미치는 영향"에 관한 연구는 대부분 다중회귀분석을 실시합니다. 독립변수가 하나 종속변수가 하나일 경우 단순회귀분석이라고 하고 독립변수가 2 개 이상이고 종속변수가 하나일 경우 다중회귀분석이라고 합니다. 독립변수는 연속 변수, 서열, 더미도 상관없지만, 종속변수는 연속변수로 이루어져야 합니다.

1) Enter(입력) 방식 다중회귀분석

1 샘플데이터	l.sav	[데이터집합6] - IBI	A SPSS Statistic	S Data Editor				
파일(F) 편	집(E) 보기(V) 데이터	(D) 변환(T)	분석(A) 다이렉트 마케팅	(<u>M</u>)	그래프(G) 유틸리티	티(U) 창(W)	도움말(<u>H</u>)
4 ; HPB29		2	C J	보고서(P) 기술통계량(E)	۴ ۴	#1 🏝 🔛		
1			노이브브마시	#	*	바다비카이 도거요ㅎ	1 <u>2</u> 21,	비스타내
		eνiπeι		명국 미파(M)	*	다 8/14	.2	1041
				일만선형보형(<u>G</u>)	×.			
1	!졸	혼자 산더	-	일만화 선형 보형(Z)	1	.00	65	중
2	1상	노인 부부만 산C	- 1	폰합 모형(<u>X</u>)	*	.00	10	하
3	졸	노인 부부만 산C	1	상관분석(<u>C</u>)	۲	.00	17	하
4	· ~	노인 부부만 산C	- 1	회귀분석(<u>R</u>)	•	🗾 자동 선형 모형화		중
5	1상	노인 부부만 산더	1	로그선형분석(<u>O</u>)	۶.	🔣 선형(L)		중
6	!졸	노인 부부만 산C	1	신경망(<u>W</u>)	۶	☑ 곡선추정(C)		중
7	:2	결혼 안한 자녀		분류분석(<u>Y</u>)	۴.	[집 이브 치스페고/이	L	중
8	!학	도인 부부만 산C	1	차원 감소(<u>D</u>)	Þ.		h	중
9	:솔 -	혼자 산다		척도(A)	F.	👪 미분형 로시스틱(<u>(G</u>)	하
10	!@	도인 부부만 산C	1	- 비모수 검정(N)		🔛 다항 로지스틱(M)	중
비 서(A)	, <u>,</u>	리기 비 서(ㅁ)	, 시청(T) i	= 크리				
군'역(A) -	• 5	411元(R) -	→ 신영(L)	로 ච덕 	T.			- W
🔚 신영 외귀문식						· 선영 회귀분석: 동계량		
			수(匝):	통계량(<u>S</u>))	회귀계수	✓ 모형 적합(M)	
NO			·기효등감	도 里(<u>)</u>		✓ 추정값(E)	📕 R 제곱 변화량(<u>S</u>)
age		-블록(B)1대상		제장(S)		🔲 신뢰구간(<u>C</u>)	📄 기술통계(<u>D</u>)	
💰 연령		이젼(⊻)		다음(N) 옵션(O)		수순(%): 95	📄 부분상관 및 편성	'관계수(P)
💰 종교		독립	변수([):			▶ 공분산 행렬(⊻)	✔ 공선성 신난(L)	
🌒 콜렌영네 🔒 학력			!양 ⋅E궤 ∧과귀			[잔차		
🚴 동거유형1						Durbin-Watson(U)		
♦ 보인부부만	산다		방법(M): (입학			📄 케이스별 진단(<u>C</u>)		e 194 - 24
✓ 결혼한사녀와산나						◎ 밖에 나타나는 미상;	(<u>(</u> 0): <u> </u> 3 ±t	6번자
▲ 동거유형2 선택변수(<u>C</u>):				규칙(U)		◎ 신제 개미소(<u>A</u>)		
응 경제상태 비 비미스 성명(C):						/्रिक् 🕴	취소 도움말	
✓ VARUUUU1 기비스 열망(C): ▲ 만성질환수 ▶								
	과거)	WLS:	가중값(H):		•	응	. 3.	에 비상 이상
♣ 건강상태2(1년전						63	12H
ſ	확?	분여넣기(P)	생설정(R) 취:	소도움말		·····································		128
	10,00					중		27#

종속변수에 "자기효능감"을 옮겨주고 독립변수에 독립변수를 넣어준 다음 "통계량"버튼 을 클릭합니다.

통계량창에서 "공정성 진단"과 "Durbin-Watson"를 체크하고 방법에 "입력"을 선택하고 확인을 클릭합니다.

모형 요약^b

모형	R	R 제곱	수정된 R 제곱	추정값의 표준오차	Durbin- Watson
1	.476 ^a	.226	.204	7.57348	1.838

a. 예측값:(상수), 운동, 영양, 스트레스관리, 대인관계

b. 종속변수: 자기효능감

분산분석ª

모형		제곱합	자유도	평균 제곱	F	유의확를
1	회귀 모형	2365.513	4	591.378	10.310	.000°
	잔차	8087.425	141	57.358		
	합계	10452.938	145			

a. 종속변수: 자기효능감

b. 예측값:(상수), 운동, 영양, 스트레스관리, 대인관계

R제곱값은 회귀모형의 설명력을 나타내는 것으로 여기서는 22.6%정도 설명력이 있다고 해석합니다.

분산분석에서 F값을 보고 회귀식이 종속변수를 설명하는 유용한지 판단합니다. 여기서 유의확률이 0.000으로 통계적으로 모형이 유의하다고 해석합니다.

Durbin-Watson값은 자기 상관을 보는 값으로 2에 가까우면 종속변수의 오차항은 자기 상관이 없이 서로 독립적이라고 판단합니다.

*	1.8	<	Durbin-Watson	1 <	2.2	\rightarrow	독립적	자기상관(오차	의	독립성이	가정됨)
---	-----	---	---------------	-----	-----	---------------	-----	---------	---	------	------

		비표준	비표준화 계수				공선성 통계량	
모형		В	표준오차	베타	t	유의확를	공차	VIF
1	(상수)	15.024	4.653		3.229	.002		
	영양	.671	.239	.229	2.810	.006	.826	1.210
	스트레스관리	.197	.234	.067	.840	.402	.856	1.168
	대인관계	.466	.188	.217	2.481	.014	.717	1.394
	운동	.643	.362	.150	1.773	.078	.763	1.310

계수ª

a. 종속변수: 자기효능감

여기서 영양(t=2.810, p<0.01)과 대인관계(t=2.481, p<0.05)만 종속변수에 통계적으로 유의미한 영향을 미쳤으며,

회귀모형은 Y(자기효능감) = 15.024 + .671(영양) + .466(대인관계)

표준화 계수 베타값은 독립변수 영향력의 상대적 크기입니다.

공차는 0.1를 초과하고 VIF는 10미만으로 모든 독립변수는 다중공정성에 문제가 없다고 해석됩니다.

2) Stepwise(단계선택) 방식 다중회귀분석

1 샘플더	0 E .sav	[데이터집합6] - IBM	SPSS Statistics	s Data Editor					
파일(F)	편집(<u>E</u>)) 보기(V) 데이터	(<u>D</u>) 변환(<u>T</u>)	분석(A) 다이렉트 마케팅	₿(<u>M</u>)	그래프(<u>G</u>) 유틸리티(U) 창(<u>W</u>)	도움말(<u>H</u>)
			5 3	보고서(<u>P)</u> 기술통계량(E)	Р р	ana 🖌	*	- C	۵ 🎹
4 : HPB2	9	2		Ŧ					
		동거유형1	노인부부만신	평균 비교(<u>M</u>)	۲	년자년와 다	동거유형2	경)	해상태
				일반선형보형(<u>G</u>)					
1	!졸	혼자 산다		일반화 선형 모형(Z)	*	.00		10	중
2	다	노인 부부만 산다	1	혼합 모형(X)	•	.00		25	하
3	· •	노인 부부만 산다	1	상관분석(<u>C</u>)	*	.00			하
4	霅	노인 부부만 산다	1	회귀분석(<u>R</u>)		📃 자동	선형 모형화		중
5	다	노인 부부만 산다	1	로그선형분석(0)		R Að	10.)		중
6	!졸	노인 부부만 산다	1	신경망(W)					중
7	·졷	결혼 안한 자녀		부름분석(Y)		2 목전	·주성(<u>C</u>)		중
8	!한	노인 부부만 산다	1	1위 가스(미)		12 일부	·최소제곱(<u>S</u>)		중
9	·졸	혼자 산다		처리고면/		🐻 이분	형 로지스틱(<u>G</u>)		하
10	!졸	노인 부부만 산다	1	역도(A)		🛄 다형	: 로지스틱(M)		중
44	本	누이 보보다 사다	1	비모수 검정(<u>N</u>)	- P				平

분석(A) → 회귀분석(R) → 선형(L)를 클릭



종속변수에 "자기효능감"을 옮겨주고 독립변수에 독립변수를 넣어준 다음 "통계량"버튼 을 클릭합니다.

통계량창에서 "공선성 진단", "Durbin-Watson", "R제곱 변화량(S)"를 체크하고 방법에 "단계선택"을 선택하고 확인을 클릭합니다.

모형 요약[°]

						į	통계량 변화량			
				추정값의					유의확를 F	Durbin-
모형	R	R 제곱	수정된 R 제곱	표준오차	R 제곱 변화량	F변화량	df1	df2	변화량	Watson
1	.382ª	.146	.140	7.87262	.146	24.655	1	144	.000	
2	.448 ^b	.201	.190	7.64203	.055	9.821	1	143	.002	1.822

a. 예측값:(상수), 대인관계

b. 예측값:(상수), 대인관계, 영양

c. 종속변수: 자기효능감

분산분석 ^a

모형		제곱함	자유도	평균 제곱	F	유의확를
1	회귀 모형	1528.095	1	1528.095	24.655	.000 ^b
	잔차	8924.843	144	61.978		
	합계	10452.938	145			
2	회귀 모형	2101.656	2	1050.828	17.993	.000°
	잔차	8351.282	143	58.401		
	합계	10452.938	145			

a. 종속변수: 자기효능감

b. 예측값:(상수), 대인관계 c. 예측값:(상수), 대인관계, 영양

Enter(입력)방법 아웃풋과 다르게 2개의 모형이 나왔습니다.

첫 번째 모형의 R제곱값은 .146, 두 번째 모형의 R제곱값은 .201로 두 번째 모형의 R제 곱값 회귀모형의 설명력이 더 컸습니다.

R제곱변화량을 살펴보면, .055가 증가 된 것을 알 수 있습니다. 물론 F 변화량 유의확률 도 .002로 통계적으로 유의했습니다.

분산분석에서 F값을 보고 회귀식이 종속변수를 설명하는 유용한지 판단합니다. 모형1과 모형2 모두 유의확률이 0.000으로 통계적으로 모형이 유의하다고 해석합니다.

계	수ª

		비표준	비표준화 계수				공선성	통계량
모형		в	표준오차	베타	t	유의확를	공차	VIF
1	(상수)	27.321	3.220		8.484	.000		
	대인관계	.821	.165	.382	4.965	.000	1.000	1.000
2	(상수)	17.311	4.469		3.873	.000		
	대인관계	.631	.172	.294	3.678	.000	.875	1.143
	영양	.734	.234	.250	3.134	.002	.875	1.143

a. 종속변수: 자기효능감

제외된 변수^a

						공선성 통계량			
모형		베타 입력	t	유의확를	편상관계수	공차	VIF	최소공차한계	
1	80 80	.250 ^b	3.134	.002	.254	.875	1.143	.875	
	스트레스관리	.147 ^b	1.863	.065	.154	.942	1.061	.942	
	운동	.173 ^b	2.022	.045	.167	.794	1.260	.794	
2	스트레스관리	.095°	1.198	.233	.100	.890	1.124	.826	
	운동	.164°	1.975	.050	.164	.793	1.261	.720	

a. 종속변수: 자기효능감

b. 모형내의 예측값:(상수), 대인관계

c. 모형내의 예측값:(상수), 대인관계, 영양

제외된 변수를 살펴보면 모형1에서 영양이 제외되었지만 유의확률이 .002로 나타나 영 양을 포함한 모형2가 구성됨.

※ 입력방식보다는 단계선택 방식의 R값이 더 작은 것은 독립변인이 줄어들었기 때문입 니다.

예제) 다중회귀분석 해석

건강증진행위가 자기효능감에 어떠한 영향을 주는지 알아보기 위해 다중회귀분석을 실시한 결과는 다음과 같다.

			Enter방	식 다중 회	귀분석		
변수	β	Std. Error	표준화 베타	t	р	공차한계	VIF
(상수)	15.024	4.653		3.229	.002		
영양	.671	.239	.229	2.810	.006	.826	1.210
스트레스관리	.197	.234	.067	.840	.402	.856	1.168
대인관계	.466	.188	.217	2.481	.014	.717	1.394
순동	.643	.362	.150	1.773	.078	.763	1.310

<표 48> 자기효능감에 미치는 영향에 대한 다중회귀분석

R²=0.226, 수정된 R²=0.204

F=10.310, p=0.000, Durbin-Watson=1.838

건강증진행위가 자기효능감에 영향을 미치는 요인을 알아보기 위하여 Enter 방식의 다중 회귀분석을 실시하였다. 다중회귀분석을 실시하기 위하여 독립변수 간의 다중공선성을 검토 하였다. 독립변수간 다중공선성은 공차한계와 VIF(분산팽창요인) 지수를 이용하였고, 독립 변수 간 VIF 지수는 1.168~1.394로 10미만이었으며, 공차한계는 0.171~0.856으로 0.1 이상으로 다중공선성이 없는 것으로 나타났다. 또한 오차의 독립성을 검증한 결과 Durbin-Watson 통계량이 1.838로 자기상관이 없는 것으로 확인되었다. 회귀분석상 모델의 설명력을 나타내는 R^2 값은 .226으로 나타나, 이 회귀모델은 건강증진행위가 자기효능감에 미치는 영향력에 대하여 약 22.6%의 설명력을 지닌다고 할 수 있다. 회귀계수의 유의성 검 정결과는 영양(t=2.810, p<.01), 대인관계(t=2.481, p<.05)가 자기효능감에 유의한 영향을 미치는 것으로 나타났다.

표준화 계수는 각 독립변인들이 종속변수인 자기효능감에 미치는 상대적인 영향력을 나타 내는 것으로 영양(β=.229)이 가장 큰 영향을 주었으며, 다음으로 대인관계(β=.217) 순이었 다. 이는 건강증진행위 중 영양과 대인관계가 좋을수록 자기효능감이 증가한다는 것을 보여 준다.

11. 더미를 활용한 다중회귀분석

차이분석결과 종속변수에 영향을 준 변인을 통제변수 또는 독립변수로 하여 다중회 귀식을 보정하려면 변인을 더미화해야 합니다.

예) 성별처럼 2개의 더미일 경우 여성=0, 남성=1

🚺 샘플데이	터.sav [데이터집합1] -	IBM SPS	S Statistics Data Editor	Contract of	the second	Same large in
파일(F) 된	편집(E) 보기(V) 데(IEI(D)	변환(T) 분석(A) 다이렉트 마케팅(M)	그래	프(G) 유틸리티(U)) 창(<u>W</u>) 도움말
			■ 변수계산(C) ☑ 케이스 내의 값 반도(O)			- A
1:성별	2		값미돌(E)			
	NO	3			종교	결혼상태
			🌆 다른 변수로 코딩변경(R)			
1	1		📷 자동 코딩변경(A)		기독교	기혼
2	2				기독교	기혼
3	3				기독교	기혼
4	4		💦 죄작의 빈 만들기(()		없음	기혼
5	5		모형화를 위한 데이터 준비(P)	•	기독교	기혼
6	6		🛃 순위변수 생성(K)		천주교	기혼
7	7		🚔 날짜 및 시간 마법사(D)		기독교	기혼
8	8				불교	기혼
9	9				기독교	사별
10	10		백합 말 뚝 값 내 제 (♥)		기독교	기혼
11	11		👹 난수 생성기(<u>G</u>)		기독교	기혼
12	12		● 변환 중지(T) Ctrl+C	3	천주교	사별
13	13		여성 68	60CH	천주교	기혼

변환(T) → 다른 변수로 코딩변경(R) 클릭



성별 선택해서 이름(N)에 "성별더미"라고 넣고 바꾸기 클릭 후 기존값 및 새로운 값(O) 클릭한다.

🖬 새로운 변수로 코딩변경: 기존값 및 새로운 값	×
- 기존값 ● 값(V): 1 ◎ 시스템-결촉값(S) ◎ 시스템 또는 사용자 결측값(U) ◎ 범위(N): 에서(T) ◎ 최저값에서 다음 값까지 범위(G): ■ 다음 강에서 친금 값까지 범위(C):	새로운 값 ④ 기준값(A): 1 ○ 시스템-결측값(Y) ④ 기존값 복사(P) 기존값> 새로운 값(D): 2> 0 추가(A) 바꾸기(C) 제거(R)
◎ 다음 값에서 최고값까지 범위(E): ◎ 기타 모든 값(0)	 ■ 출력변수가 문자열임(B) 너비(W): 8 ■ 숫자형 문자를 숫자로 변환(5->5)(M) 계속 취소 도움말

남성이 1, 여성이 2로 코딩되어 있고, 여성을 기준으로 바꾸려면, 여성 2를 0으로 바꾸 고 남성은 그대로 1로 바꿔줍니다.

🔄 *샘플데이티	H.sav [데이터집합1] -	IBM SPSS Statistics	Data Editor						
파일(F) 편집	집(E) 보기(V) 데이	비터(<u>D</u>) 변환(<u>T</u>) -	분석(A) 다이렉트 마케팅(M)) 그래프(<u>G</u>)	유틸리티(<u>U</u>) :	창(<u>W</u>) 도움말(<u>H</u>)			
🔁 H		r 🧃 🛿	ă 📥 🚅 🃭	H 👪		▲ 🔳 .			
	NO	성별	성별더미	age	연령	종교			
1	1	여성	_00	65	60CH	기독교			
2	2	남성	1.00	78	70 CH	기독교			
3	3	남성	1.00	84	<mark>80</mark> 대 미상	기독교			
4	4	며성	.00	73	70 CH	없음			
5	5	남성	1.00	70	70CH	기독교			
6	6	며성	.00	65	60CH	천주교			
7	7	여성	.00	65	60CH	기독교			
8	8	남성	1.00	65	60CH	불교			
9	9	며성	.00	70	70CH	기독교			
10	10	며성	_00	73	7 <mark>0</mark> CH	기독교			
11	11	며성	.00	75	70CH	기독교			
12	12	남성	1.00	68	60CH	천주교			
13	13	여성	.00	68	60CH	천주교			
14	14	남성	1.00	71	70CH	기독교			
15	15	며성	.00	75	<mark>70</mark> 대	불고			
16	16	남성	1.00	73	700	없음			

여성이 0이고, 남성이 1인 성별더미가 생성됩니다.

회귀식을 돌릴 경우 성별더미를 독립변수로 넣고 돌리면 됩니다.



이번엔 연령을 더미를 만들기 위해, 우선 연령을 "연령더미1" 넣고 바꾸기(H) 후 기존 값 및 새로운 값(O)를 누루고 "60대" 1, "70대" 2, "80대 이상" 3으로 코딩되어 있는 것을 1을→0으로, 2를→1로, 3을→0으로 바꾸어줍니다.



두 번째로 "연령더미2"를 만들어 줍니다. 연령더미2는 1을→0으로, 2를→0으로, 3을→1 로 바꾸어 줍니다.

*샘플	레이터.sav [데이터집합1] - IB	M SPSS Statistics Data Edit	or			1 1 1 1 m		
파일(E)	편집(E)	보기(V) 데이트	H(D) 변환(T) 분석(A)	다이렉트 마케팅(<u>M</u>)	그래프(G)	유틸리티(U) 창(W) 도	움말(<u>H</u>)		
) 🛄 🖬	r 🤉 🖪 🛓				I 🗛 🔷 🌑	46	
1:연령_더미1 .0 표시									
		성별	성별더미	age	연령	연령더미1	연령더미2	종교	
1	1	여:	섬 .00	65	60CH	.00	.00	기독교 🖆	
2	2	남:	섬 1.00	78	70CH	1.00	.00	기독교	
3	3	남:	섬 1.00	84	80대 이상	.00	1.00	기독교	
4	4	여:	섬 .00	73	70CH	1.00	.00	없음	
5	5	남:	섬 1.00	70	70CH	1.00	.00	기독교	
6	6	04:	섬 .00	65	60CH	.00	.00	천주교	
7	7	여	섬 .00	65	60CH	.00	.00	기독교	
8	8	남:	섬 1.00	65	60CH	.00	.00	불교	
9	9	여:	섬 .00	70	70CH	1.00	00.	기독교	
10	10	04:	섬 .00	73	70CH	1.00	.00	기독교	
11	11	04:	섬 .00	75	70CH	1.00	.00	기독교	
12	12	남	섬 1.00	68	60CH	.00	00_	천주교	
13	13	여:	섬 .00	68	60CH	.00	.00	천주교	
14	14	남	섬 1.00	71	70CH	1.00	00.	기독교	
15	15	여	섬 .00	75	700	1.00	.00	불교	
16	16	남	섬 1.00	73	700	1.00	.00	없음	
17	17	0i:	섬 .00	65	60CH	.00	.00	없음	
데이터	데이터 보기(0) 변수 보기(0)								

그 결과는 60대는 연령더미1과 연령더미2가 00으로, 70대는 10으로, 80대 이상은 01로 되었습니다. 여기서 중요한 것은 기준이 되는 값을 0으로 만들어주는 것입니다.

다중회귀식을 돌릴 때는 "연령더미1"과 "연령더미2"를 함께 넣고 돌립니다.



위는 Syntax(명령어)로 더미를 쉽게 만드는 법입니다. 연령=1 일 때 연령_더미1=0으로, 연령_더미2=0으로 만든다. 연령=2 일 때 연령_더미1=1로, 연령_더미2=0으로 만든다. 연령=3 일 때 연령_더미1=0으로, 연령_더미2=1로 만든다. 만약 연령이 4가지 범주로 되어 있다면,

🔚 *4개_더미.sps - IBM SPSS Statisti	cs Syntax Editor	-	-	-maxi			
파일(F) 편집(E) 보기(V) 데이	터(D) 변환(T)	분석(<u>A</u>)	다이렉트 마케팅(M) 그래프(G)	유틸리티(U)	실행(<u>R</u>)	도
							(
	>∄ ♦/∄ ।			II	· 활성: 데이E	러집합1 ▼	
if	1 if(연	령 =1) 연령	킹 더미1=0.				
if	2	ć.	100				
if F	3 if(연*	령 =1) 연령	령_더미2=0.				
if	4 5 if(연*	텸 =1) 연령	킹 (카미)3=0				
if	6	0 ./ 20					
if	7 ▶ if(연	텸 =2) 연령	\$_더미1=1.				
if	8 9 ; f/01:	려	f [][]2=0				
if	10	6 -2) C 6	5_01012-0.				
if	11 if(연*	령 =2) 연령	령_더미3=0.				
if	12	51010-15					
execute.	13 II(원)	8 = 3) 전망	3_[]0]1=0.				
	15 if(연	텸 =3) 연령	킹_더미2= <mark>1</mark> .				
	16						
	17 if(연· 18	명 =3) 연당	호_더미3=0.				
	19 if(연*	령 =4) 연령	킹 더미1=0.				
	20						
	21 if(연 ³	령 =4) 연령	령_더미2=0.				
	22 23 if(연3	텸 =4) 연종	5 CHO13=1.				
	24	,	10001 10				
	25 exec	cute.					
	26						

위와 같이 연령의 범주가 4개일 때

연령=1 일 때 연령_더미1=0으로, 연령_더미2=0, 연령_더미3=0으로 만든다. 연령=2 일 때 연령_더미1=1로, 연령_더미2=0으로 연령_더미3=0으로 만든다. 연령=3 일 때 연령_더미1=0으로, 연령_더미2=1로 연령_더미3=0으로 만든다. 연령=4 일 때 연령_더미1=0으로, 연령_더미2=0로 연령_더미3=1으로 만든다.

그럼 연령 1이 기준인 0으로 됩니다.

1 *샘플더	이터.sav [데	이터집합1] - IBM :	SPSS Statistic	cs Data Editor				
파일(F)	편집(<u>E</u>) 보	보기(V) 데이터(D) 변환(<u>T</u>)	분석(A) 다이렉트 마케팅	9 (W)	그래프(<u>G</u>)	유틸리티(U) 창(W) 도·	움말(<u>H</u>)
			7	보고서(<u>P)</u> 기술통계량(E)	*	11 🍇	🖬 🔤 🐴 🛛	
1:연령_[301	.0		Ŧ		1		
		성별	성별(평균 비교(<u>M</u>)	۲	연령	연령더미1	연령더미2
1	1	며성		일반선형모형(<u>G</u>)	۲	60CH	.00	.00
2	2	남성		일반화 선형 모형(Z)	•	70CH	1.00	.00
3	3	남성		혼합 모형(X)	۲	80대 이상	.00	1.00
4	4	며성		상관분석(<u>C</u>)	*	70CH	1.00	.00
5	5	남성		회귀분석(<u>R</u>)	F.	🗾 자동 선형	모형화 00	.00
6	6	며성		로그선형분석(<u>O</u>)	•	🚮 선형(L)	00	.00
7	7	며성		신경망(<u>W</u>)	•	☞ 곡선추정(C)	. <mark>00</mark> .
8	8	남성		분류분석(Y)		이 비 귀 사	00	.00
9	9	며성		차원 감소(<u>D</u>)	•	11월 월두 쇠오.	AII TE (2)	. <mark>00</mark> .
10	10	며성		적도(A)	*	👪 이분형 로	지스틱(<u>G</u>) 00	.00
11	11	며성		 비모수 검정(N)		🔛 다항 로지	스틱(<u>M</u>) 00	.00

이제 만들어진 더미를 가지고 다중회귀분석을 실시하겠습니다. 분석(A) → 회귀분석(R) → 선형(L)을 클릭

종속변수에 "자기효능감"을 성별과 연령에 따른 차이가 나타나 성별과 연령을 더미로 만들어서 독립변수로 넣고, 이전의 독립변수인 건강증진행위 하위요인을 독립변수로 놓고 "통계량(S)"를 클릭 후 "공선성 진단(L)"과 "Durbin-Watson(U)"를 선택하고 회귀식을 돌립니다.

계수ª	
-----	--

		비표준화 계수		표준화 계수			공선성	통계량
모형		В	표준오차	베타	t	유의확를	공차	VIF
1	(상수)	21.628	4.743		4.560	.000		
	성별더미	4.007	1.371	.206	2.923	.004	.929	1.076
	연령더미1	-2.872	1.648	170	-1.743	.084	.487	2.054
	연령더미2	-6.571	1.817	365	-3.616	.000	.453	2.206
	영양	.601	.221	.205	2.723	.007	.814	1.229
	스트레스관리	.114	.216	.039	.528	.599	.847	1.181
	대인관계	.442	.173	.206	2,556	.012	.711	1.407
	운동	.456	.335	.107	1.360	.176	.751	1.331

a. 종속변수: 자기효능감

더 많은 아웃풋이 있지만, 앞에서 다루었으니 그 부분을 참고 하시고 여기서는 실질적으 로 어떻게 더미독립변수가 해석되어지는지를 살펴보겠습니다.

성별더미와 연령더미2가 통계적으로 유의미하게 나왔습니다 성별더미는 "남성"입니다. 이유는 "여성"이 0이기 때문입니다. 그러면 연령은 "60대"가 0이니 연령더미1은? "70 대"이고, 연령더미2는 "80대 이상"이 되겠죠?

이를 해석하면, 성별이 남성일수록 자기효능감이 증가하였으며, 연령이 80대 이상일 때 60대보다 자기효능감이 감소하는 것으로 나타났다. 이렇게 해석을 하시면 됩니다.

더미 만들기 더	병령어(Syntax)
1) 3개 더미 만들기	
	3) 5개 더미 만들기
if(연령 =1) 연령_더미1=0.	
if(연령 =1) 연령_더미2=0.	if(연령 =1) 더미1=0.
	if(연령 =1) 더미2=0.
if(연령 =2) 연령_더미1=1.	if(연령 =1) 더미3=0.
if(연령 =2) 연령_더미2=0.	if(연령 =1) 더미4=0.
if(여려 =3) 여려 더미1=0	if(연령 =2) 더미1=1.
if(연령 =3) 연령 더미2=1	if(연령 =2) 더미2=0.
	if(연령 =2) 더미3=0.
execute.	if(연령 =2) 더미4=0.
2) 4개 더미 만들기	if(연령 =3) 더미1=0.
	if(연령 =3) 더미2=1.
if(연령 =1) 연령_더미1=0.	if(연령 =3) 더미3=0.
if(연령 =1) 연령_더미2=0.	if(연령 =3) 더미4=0.
if(연령 =1) 연령_더미3=0.	
	if(연령 =4) 더미1=0.
if(연령 =2) 연령_더미1=1.	if(연령 =4) 더미2=0.
if(연령 =2) 연령_더미2=().	if(연령 =4) 더미3=1.
if(연령 =2) 연령_더미3=0.	if(연령 =4) 더미4=0.
if(연령 =3) 연령_더미1=0.	if(여려 =5) 더미1=0
if(연령 =3) 연령_더미2=1.	if(여려 =5) 더미2=0
if(연령 =3) 연령_더미3=0.	if(여려 =5) 더미3=0
	if(연령 =5) 더미4=1.
if(연령 =4) 연령_더미1=0.	
if(연령 =4) 연령_더미2=0.	execute.
if(연령 =4) 연령_더미3=1.	
execute.	

여기서 "연령"을 바꾸고 싶은 변수명으로 바꾸고 나서 명령어(Syntax)창에 복사해 서 사용하시면 편합니다.

예제) 더미변수를 활용한 회귀분석 해석

응답자의 인구학적 특성, 건강증진행위가 자기효능감에 미치는 영향을 검토하기 위해 더미 변수를 활용한 다중 회귀 분석을 실시한 결과는 다음과 같다.

		Enter방식 다중 회귀분석									
	변수	β	Std. Error	표준화 베타	t	р	공차한계	VIF			
	(상수)	21.628	4.743		4.560	.000					
인구	성별더미(남성)	4.007	1.371	.206	2.923	.004	.929	1.076			
학적	연령더미1(70대)	-2.872	1.648	170	-1.743	.084	.487	2.054			
변인	연령더미2(80대 이상)	-6.571	1.817	365	-3.616	.000	.453	2.206			
	영양	.601	.221	.205	2.723	.007	.814	1.229			
건강 즈기	스트레스관리	.114	.216	.039	.528	.599	.847	1.181			
등신 행위	대인관계	.442	.173	.206	2.556	.012	.711	1.407			
	운동	.456	.335	.107	1.360	.176	.751	1.331			

<표 56> 자기효능감에 미치는 영향에 대한 다중회귀분석

R²=0.363, 수정된 R²=0.331

F=11.244, p=0.000, Durbin-Watson=1.892

응답자의 인구학적 특성(성별, 연령)과 건강증진행위가 자기효능감에 영향을 미치는 요인 을 알아보기 위하여 Enter 방식의 다중회귀분석을 실시하였다. 다중회귀분석을 실시하기 위 하여 독립변수 간의 다중공선성을 검토하였다. 독립변수간 다중공선성은 공차한계와 VIF(분 산팽창요인) 지수를 이용하였고, 독립변수 간 VIF 지수는 1.076~2.206으로 10미만이었으 며, 공차한계는 0.453~0.929로 0.1 이상으로 다중공선성이 없는 것으로 나타났다. 또한 오 차의 독립성을 검증한 결과 Durbin-Watson 통계량이 1.892로 자기상관이 없는 것으로 확 인되었다. 회귀분석상 모델의 설명력을 나타내는 R^2 값은 .363으로 나타나, 이 회귀모델은 인구학적 특성과 건강증진행위가 자기효능감에 미치는 영향력에 대하여 약 36.3%의 설명력 을 지닌다고 할 수 있다. 회귀계수의 유의성 검정결과는 성별더미(t=2.923, p<.01), 연령더 미2(t=-3.616, p<.001), 영양(t=2.723, p<.01), 대인관계(t=2.556, p<.05)가 자기효능감에 유의한 영향을 미치는 것으로 나타났다.

표준화 계수는 각 독립변인들이 종속변수인 자기효능감에 미치는 상대적인 영향력을 나타 내는 것으로 연령더미2(β=-.365)가 가장 큰 영향을 주었으며, 다음으로 성별더미와 대인관 계(β=.206), 영양(β=.205) 순이었다. 이는 성별이 남성이고 건강증진행위 중 영양과 대인관 계가 좋을수록 자기효능감이 증가하며, 연령이 80대 이상일 때 자기효능감이 감소하다는 것 을 보여준다.

Ⅲ. 중급 통계 분석

중급 통계 분석에서는 위계적 다중회귀분석을 활용한 조절효과분석, 매개효과분석, 이항로지스틱회귀분석, 반복측정분산분석, 공분산분석, 정규성, 비모수 통계를 다루 도록 하겠습니다.

1. 위계적 다중회귀분석

다중회귀분석은 독립변수들을 동시에 넣고 돌리기 때문에 독립변수들의 변화량을 쉽 게 알 수는 없지만, 위계적 다중회귀분석은 변수들을 순차적으로 넣을 수 있기 때문에 단계별 차이와 설명력을 알 수 있습니다. 이를 통해 조절효과분석도 가능합니다.

1 샘플데이터.	sav [데이터집합1] -	IBM SPSS Statistics	Data Editor							
파일(F) 편집	¦(E) 보기(⊻) 데	미터(D) 변환(T)	분석(A) 다이렉트 마케팅(M)	그래프(G)	유틸리티(U)	창(<u>W</u>) 도움말(H)				
		5 3	보고서(P) ▶ 기술통계량(E) ▶	#		- 🐴 🎹 🛃				
			∄ ►		6-0-	<i>y</i>				
	NO	성별	평균비교(<u>M</u>) ▶	age	연령	연령더미1				
4	4	며실	일반선형보형(<u>G</u>) 🕨	73	700	1.00				
5	5	남성	일만화 선형 보형(∠) ▶	70	700	1.00				
6	6	며성	혼합 모형(X) ▶	65	60CH	.00				
7	7	며성	상관분석(<u>C</u>) ▶	65	60CH	.00				
8	8	남성	회귀분석(民) ▶	📕 자동 선형	모형화	.00				
9	9	며실	로그선형분석(⊙) ▶	🔝 선형(L)		1.00				
10	10	여상	신경망(₩) ▶	🗾 곡선추정	(C)	1.00				
11	11	여성	분류분석(Y) ▶	國 익보 치소	· 제공(S)	1.00				
12	12	남성	차원 감소(<u>D</u>) ▶			.00				
13	13	여실	척도(<u>A</u>) ▶	📷 미운영 도	지스틱(G)	.00				
14	14	남성	비모수 겸정(ℕ) ▶		스틱(<u>M</u>)	1.00				
15	15	며실	예측(T) ▶	🔛 순서(<u>D</u>)	2	1.00				
16	16	남성	새조하류/이 ⊾	🔛 프로빗(P)	1.00				
분석(A) →	→ 회귀분석(F	X) → 선형(L	,)을 클릭							
🚺 선형 회귀분석	14.00	-	22) 🕼 선형 회국	귀분석: 통계량	×				
▲ 선황 최귀분석 ▲ 선황 최귀분석: 통계량 ▲ ▲ 선황 최귀분석: 통계량 ▲ 선황 최귀분석: 통계량 ▲ ▲ 바PB32 ▲ 차기효능감 도표[D ▲ 바PB33 ▲ 바PB33 ▲ 北기효능감 도표[D ▲ 바PB34 ● 비PB34 ● 보상관 및 편상관계수(P) ● 바PB34 ● 비PB35 ● 로변수(D): ▲ 전성 진단(L) ● 조리스관리 ● 대일관계 ● 물분상관 및 편상관계수(P) 용건성 진단(L) ● 소트레스관리 ● 레인관계 ● 물변 수(D): ● 물변 (P): ● 소트레스관리 ● 레인(P): ● 물변 (P): ● 물변 (P): ● 조리스 관리 ● 레인(P): ● 물변 (P): ● 물변 (P): ● 조건성 진단(L) ● 출연 (P): ● 출연 (P): ● 출연 (P): ● 지수 10 스 별 72*C ● 출연 (P): ● 출연 (P): ● 출연 (P): ● 지수 10 스 별 72*C ● 출연 (P): ● 출연 (P): ● 출연 (P): ● 전성 진단(C): ● 출연 (P): ● 출연 (P): ● 취례 LEEL + E 이상값(O): 3 ● 표준 편차 ● 전 연 (P): ● 출연 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): ● 전 (P): </td										

종속변수를 넣고 독립변수에 인구학적 특성인 age를 넣습니다.

통계량(S)를 클릭하고 "R제곱 변화량(S)", "공선성 진단(L)", "Durbin-Watson"를 체크하 고 계속을 클릭한 후 다음 버튼을 누릅니다.





다음 버튼을 클릭하고 독립변수인 "영양"을 넣어 주고 다시 "다음(N)"을 클릭합니다.

마지막으로 독립변수에 "대인관계"를 넣고 확인을 누릅니다. 확인을 누르기 전에 통계량(S)를 클릭해서 앞에서 한 작업을 하셔도 상관은 없습니다.

모형 요약^d

						통계량 변화량							
				추정값의					유의확를 F	Durbin-			
모형	R	R 제곱	수정된 R 제곱	표준오차	R 제곱 변화량	F변화량	df1	df2	변화량	Watson			
1	.449 ^a	.202	.196	7.61272	.202	36.368	1	144	.000				
2	.522 ^b	.273	.262	7.29153	.071	13.965	1	143	.000				
3	.569°	.324	.310	7.05513	.051	10.744	1	142	.001	2.000			

a. 예측값:(상수), age

b. 예측값:(상수), age, 영양

c. 예측값:(상수), age, 영양, 대인관계

d. 종속변수: 자기효능감

			분산분석"								
모형		제곱함	자유도	평균 제곱	F	유의확를					
1	회귀 모형	2107.640	1	2107.640	36.368	.000 ^b					
	잔차	8345.298	144	57.953							
	합계	10452.938	145								
2	회귀 모형	2850.132	2	1425.066	26.804	.000°					
	잔차	7602.806	143	53.166							
	합계	10452.938	145								
3	회귀 모형	3384.916	3	1128.305	22.668	.000 ^d					
	잔차	7068.022	142	49.775							
함계 10452.938 145											
a. 2	5속변수:자기	효능감									

b. 예측값:(상수), age c. 예측값:(상수), age, 영양

t. 에득畝. (공구), age, 영양, 대인관계 d. 예측값: (상수), age, 영양, 대인관계

그러면 3가지 모델 아웃풋이 나옵니다.

모형 1은 인구학적 특성인 "age"만 넣은 모형이고, 모형 2는 인구한적 특성 + 독립변수 "영양"을 넣은 모형 모형 3은 인구한적 특성 + 독립변수 "영양" + 독립변수 "대인관계"을 넣은 모형입니다.

R제곱 변화량과 F변화량 그리고 F변화량 유의확률을 보면 모형1에서 독립변수 "영양" 이 투입되었을 때 F변화량 유의확률이 .000으로 모형1보다 모형2가 R제곱이 .071이 증 가하였으며, 이는 통계적으로 유의미하다. 만약 F변화량 유의확률이 .05보다 컸다면, 모 형1에서 투입된 독립변수 "영향"은 R제곱값이 증가하지 않았다는 뜻입니다. 이 부분은 다음에 조절효과에서 중요하게 해석이 됩니다.

모형1, 모형2, 모형3의 분산분석 적합도는 모두 유의하였습니다.

계수ª

		비표준:	화 계수	표준화 계수			공선성	통계량
모형		В	표준오차	베타	t	유의확를	공차	VIF
1	(상수)	85.220	7.033		12.118	.000		
	age	554	.092	449	-6.031	.000	1.000	1.000
2	(상수)	65.040	8.633		7.534	.000		
	age	484	.090	392	-5.379	.000	.957	1.045
	영양	.798	.214	.273	3.737	.000	.957	1.045
3	(상수)	56.494	8.751		6.456	.000		
	age	446	.088	361	-5.078	.000	.940	1.064
	영양	.565	.219	.193	2.582	.011	.855	1.170
	대인관계	.524	.160	.244	3.278	.001	.860	1.163

a. 종속변수: 자기효능감

해석은 모형3을 기준으로 해석합니다. 위계적 회귀분석은 독립변수들이 순차적으로 들 어갔을 때의 변화량을 아는 것이 메인입니다.

2. 위계적 회귀분석을 활용한 조절효과분석

조절변수가 종속변수와 독립변수 관계를 조절한다는 것을 알아보기 위해 위계적 회 귀분석을 활용한 조절효과분석을 실시합니다.

🚺 위계적	🔒 위계적_조절효과:sav [데이터집합2] - IBM SPSS Statistics Data Editor												
파일(F)	편집(E) 보기(V)	데이터(D)	변환(T)	분석(<u>A</u>)	다이렉트 마케팅(M)	그래	표(<u>G</u>)	유틸리티	(<u>U</u>)	황(<u>W</u>)	도움말(난		
			📄 변수 🔀 케이:	<mark>계산(C)</mark> 스 내의 값	(빈도(<u>O</u>)		*,			5			
		~~~	값이	동(E)						- 70			
	ID	면영	7+0	ш			독립면	Ŷ		조설턴	1 <del>2</del>		
1	1	400	[ [ ] 같는	면수도 꼬	· 3면경( <u>5</u> )			2.9	0		3.56		
2	2	400	🔤 다른	변수로 코	.딩변경( <u>R</u> )			2.4	7		3.48		
3	3	200	🔣 자동	코딩변경	( <u>A</u> )			1.7	3		3.85		
4	4	200	▶ 비주 위	걸 빈 만들	71(B)			2.3	7		3.15		
5	5	200	₩ 최적의	의 빈 만들	70			2.2	3		4.00		
6	6	200	모행	하루 이하	데이터 주비(p)			3.3	0		3.52		
7	7	200		이글 카인		-		4.3	7		4.85		
8	8	200	1 순위	<u> </u>	( <u>K</u> )	_		3.1	3		2.67		
9	9	200	🗎 날짜	및 시간 🛛	바법사( <u>D</u> )			2.0	3		3.59		

조절효과를 보려면 먼저 상호작용항을 만들어야 합니다. 변수계산(C)를 클릭



독립변수X조절변수= 상호작용항을 계산해 줍니다. 계산은 *(곱하기)로 해주면 됩니다.

🔚 পাৰ	a 위계적_조절효과.sav [데이터집합2] - IBM SPSS Statistics Data Editor												
파일(F)	편집( <u>E</u> )	보기(V)	데이터(D) 부	변환(T)	분석(A) 다이렉트 마케팅	( <u>M</u> )	그래프(G) 유틸리티(U	<u>)</u> 창( <u>W</u> )	도움말(H	)			
				2	보고서(P) 기술통계량(E)	•	M 🕷 🖬	<b>-</b>		A 0 6			
					Ŧ	•							
		ID	연령	연령?	평균 비교(M)	*	독립변수	조절변	!수	독립변수X조절변수			
1		1	40CH	300	일반선형모형(G)		2.90		3.56	10.31			
2		2	40CH	300	으 일반화 선현 모현(Z)		2.47		3.48	8.59			
3		3	2008	200	<u>さむ口 さい</u>		1.73		3.85	6.68			
4		4	2008	200			2.37		3.15	7.45			
5		5	2008	200	상관문적(단)	•	2.23		4.00	8.93			
6		6	2008	200	회귀분석( <u>R</u> )	ħ.,	屋 자동 선형 모형화		3.52	11.61			
7		7	20CH	200	로그선형분석( <u>O</u> )	*	🚮 선형(L)		4.85	21.19			
8		8	20CH	200	신경망( <u>W</u> )	۲	📝 곡선추정(C)		2.67	8.36			
9		9	20CH	200	분류분석( <u>Y</u> )	۲	N 인터 치스페고(Q)		3.59	7.30			
10		10	20CH	200	차원 감소( <u>D</u> )	•	·····································		4.41	8.23			
11		11	20CH	200	쳑도(A)		₩ 미분형 로시스틱( <u>G</u> ).		3.85	5.91			
12		12	2008	200	_ 비모수 검정(N)		🔛 다항 로지스틱( <u>M</u> )		2.70	9.46			
13		13	2008	200	(山本(工)		🔣 순서(D)		3.81	9.41			
14		14	20CH	200	생존확률( <u>S</u> )	•	🔛 프로빗(巴)		3.52	7.51			

빨간 사각형을 보시면 상호작용항인 "독립변수X 조절변수"가 생성된 걸 확인할 수 있 습니다.

위계적 다중회귀를 하기 위해 분석(A) → 회귀분석(R) → 선형(L)을 클릭합니다.

🕼 선형 회귀분석		×
<ul> <li>✔ ID</li> <li>✔ 연령</li> <li>✔ 연령집단</li> <li>✔ 독립변수</li> <li>✔ 조절변수</li> <li>✔ 독립변수X조절변수</li> </ul>	종속변수(D):	통계량(S) 도표(T) 저장(S) 옵션(Q) 붓스트랩(B)
	선택변수( <u>C</u> ): 규착( <u>U</u> ) 케이스 설명( <u>C</u> ): WLS 가중값( <u>H</u> ):	
확인	붙여넣기(P) 재설정(R) 취소 도움말	

앞에서 위계적 회귀분석을 실시한 것과 같이 종속변수와 독립변수를 넣고 "다음(N)"을 클릭합니다. 여기서는 인구학적 특성 "연령"을 넣어줍니다.

🕼 선형 회귀분석			X
<ul> <li>✔ ID</li> <li>✔ 연령</li> <li>✔ 연령집단</li> <li>✔ 독립변수</li> <li>✔ 조절변수</li> <li>✔ 독립변수×</li> </ul>	종속변수(D):	[다음N) 입력 규칙(U)	통계량( <u>S</u> ) 도표( <u>T</u> ) 저장(S) 옵션(Q) 붓스트랩(B)
"독립변수"를 넣고 "다음(N)'	"을 클릭		
대 신영 회귀문석	종속변수( <u>D</u> ): · · · · · · · · · · · · · · · · · · · · · · · · · · · · · ·	[다음N) (다음N) 입력 🔹	통계량(S) 도표(T) 저장(S) 옵션(Q) 봇스트랩(B)
"조절변수"를 넣고 "다음(N)" ▲ 선형 회귀분석 《 10 ④ 연령 ④ 연령 ● 역량 ● 역명 ● () () () () () () () () () () () () ()	" 을 클릭 값 통계량(S), 도표(T) 저장(S) 유선(Q) 용스트랩(B) 양법(M): 입력 : : : : : : : : : : : : :	<ul> <li>선형 회공분석: 통계량</li> <li>회귀계수</li> <li>초정값(E)</li> <li>신뢰구간(C)</li> <li>수준(%): 95</li> <li>공분산 행렬(V)</li> <li>· · · · · · · · · · · · · · · · · · ·</li></ul>	<ul> <li>✓ 모형 적합(M)</li> <li>✓ R 제곱 변화량(S)</li> <li>기술통계(D)</li> <li>부분상관 및 편상관계수(P)</li> <li>✓ 공선성 진단(L)</li> <li>값(Q): 3 표준편차</li> <li>최소 도움말</li> </ul>

마지막으로 상호작용항을 넣고 통계량(S)에서 "R제곱 변화량(S)", "공선성 진단(L)", "Durbin-Watson"를 체크하고 계속을 클릭한 후 다음 버튼을 클릭합니다.

모형 요약^e

						통계량 변화량						
모형	R	R제곱	수정된 R 제곱	추정값의 표준오차	R 제곱 변화량	F번화량	df1	df2	유의확를 F 변화량	Durbin- Watson		
1	.002ª	.000	003	.32091	.000	.002	1	359	.969			
2	.573 ^b	.329	.325	.26327	.329	175.392	1	358	.000			
3	.635°	.403	.398	.24854	.075	44.683	1	357	.000			
4	.640 ^d	.410	.403	.24753	.007	3.933	1	356	.048	2.012		

a. 예측값:(상수), 연령

b. 예측값:(상수), 연령, 독립변수

c. 예측값:(상수), 연령, 독립변수, 조절변수

d. 예측값:(상수), 연령, 독립변수, 조절변수, 독립변수X조절변수

e. 종속변수: 종속변수

분산분석^a

모형		제곱함	자유도	평균 제곱	F	유의확를
1	회귀 모형	.000	1	.000	.002	.969 ^b
	잔차	36.971	359	.103		
	합계	36.971	360			
2	회귀 모형	12.157	2	6.078	87.697	.000°
	잔차	24.814	358	.069		
	합계	36.971	360			
3	회귀 모형	14.917	3	4.972	80.493	.000 ^d
	잔차	22.054	357	.062		
	합계	36.971	360			
4	회귀 모형	15.158	4	3.790	61.849	.000 ^e
	잔차	21.813	356	.061		
	합계	36.971	360			

a. 종속변수: 종속변수

b. 예측값:(상수), 연령

c. 예측값:(상수), 연령, 독립변수

d. 예측값: (상수), 연령, 독립변수, 조절변수 e. 예측값: (상수), 연령, 독립변수, 조절변수, 독립변수X조절변수

아웃풋을 살펴보면, 총 4개의 모형이 있습니다. 마지막으로 상호작용항을 넣은 모형이 모형4군요. 여기서 중요한 것은 통제변수를 제외한 독립변수, 조절변수, 상호작용항의 F

변화량 유의확률입니다. 모두 0.05보다 작아 통계적으로 유의미하게 나왔습니다. 이는 종속변수와 독립변수 간의 영향관계를 조절변수가 조절한다고 해석할 수 있습니 다.

	계수 ^a										
		비표준	화 계수	표준화 계수			공선성	통계량			
모형		В	표준오차	베타	t	유의확를	공차	VIF			
1	(상수)	2.050	.048		42.455	.000					
	연령	.001	.018	.002	.039	.969	1.000	1.000			
2	(상수)	1.247	.072		17.221	.000					
	연령	.032	.015	.096	2.192	.029	.974	1.027			
	독립변수	.285	.022	.581	13.244	.000	.974	1.027			
3	(상수)	1.973	.128		15.379	.000					
	연령	.055	.014	.164	3.841	.000	.919	1.088			
	독립변수	.265	.021	.541	12.922	.000	.954	1.049			
	조절변수	202	.030	286	-6.684	.000	.913	1.095			
4	(상수)	1.155	.432		2.675	.008					
	연령	.054	.014	.161	3.799	.000	.918	1.089			
	독립변수	.581	.161	1.185	3.618	.000	.015	64.772			
	조절변수	.022	.117	.032	.192	.848	.060	16.579			
	독립변수X조절변수	087	.044	671	-1.983	.048	.014	68.983			

a. 종속변수: 종속변수

상호작용항의 유의확률이 0.048로 유의수준 0.05보다 작아 통계적으로 유의미했습니다. 여기서 공차와 VIF는 상호작용항이 들어갔기 때문에 범위를 초과했네요. 별의미는 없 습니다.

### 예제) 위계적 회귀분석을 활용한 조절효과분석

조절변수의 조절효과를 검증하기 위하여 연령을 통제한 후 조절변수의 상호작용항을 구성 하여 위계적 회귀분석을 실시하였다. 1단계에선 통제변수를, 2단계에선 독립변수, 3단계에 선 조절변수를, 4단계에선 독립변수와 조절변수를 조합한 상호작용항을 투입하여 분석하였 다.

단형 제1모형 제2모형 제3모형

<표 64> 종속변수와 독립변수안의 관계에서 조절변수의 조절효과

	모형	제11	모형	제	2모형	제	3모형	제	4모형
변수	구분	eta	t	$\beta$	t	$\beta$	t	eta	t
통제 <u>변수</u>	연령	.002	.039	.096	$2.192^{*}$	.164	3.841***	.161	3.799***
	독립변수			.581	13.244***	.541	12.922***	1.185	3.618***
	조절변수					286	-6.684***	.032	.192
	독립변수X조절	변수						671	-1.983*
	$R^2$	.00	)2		329	•	403	.410	
	수정된 $R^2$	0	03		325		398	.403	
	$R^2$ 변화량	.00	00	.329 .075		.007			
	F 변화량	.00	)2	175	.392***	44.683***		3.933*	
	F	.00	)2	87.	697***	80.493***		61.849***	

*p< .05, ***p< .001

종속변수와 독립변수의 관계에서 조절변수의 조절효과를 살펴본 결과 조절변수는 조절효 과가 있었다. 이를 자세히 살펴보면 다음과 같다. 독립변수가 종속변수에 미치는 영향과 조 절변수의 조절효과를 살펴보기 위해 제1모형에서는 통제변수인 연령의 종속변수에 대한 영 향력을 살폈다. 그 결과 모델의 적합성은 F=.002, p>.05로 나타나 통계적으로 유의하지 않 았다.

제 2모형은 연령을 통제한 상태에서 독립변수의 영향력을 알아본 결과이다. 그 결과 모형 1에 비해 종속변수에 미치는 영향력이 유의수준 p<.001수준에서 32.9%증가한 것으로 나타 났다. 변수별 영향력에서도 종속변수에 독립변수(t=13.244, p<.001)이 영향을 미치고 있었 다. 이는 독립변수가 증가할수록 종속변수가가 증가한다는 것을 의미한다.

제3모형은 조절변수가 종속변수와 독립변수의 관계에서 조절효과를 가질 수 있는 지를 알아보기 위한 단계이다. 분석결과 조절변수가 투입된 모형 3의 적합성은 통계적으로 유의 하였으며(F=80.493, p<.001), 이전 단계에 비해 모형의 설명력이 7.5%(p<.001) 증가하여 조절변수가 종속변수에 미치는 영향력이 있다는 것과 조절변수의 조절효과가 있음을 보여주 고 있다. 또한 종속변수와 조절변수의 관계(β=-.286)는 부(-)의 영향관계가 있는 것으로 나 타났다. 이는 곧 조절변수가 증가할수록 종속변수는 감소한다는 것을 의미한다. 결국 모형 3에서 보여준 결과는 조절변수가 종속변수에 영향을 미칠 뿐만 아니라 조절효과를 가질 수 있음을 보여준 것이다.

제 4모형은 조절변수가 종속변수와 독립변수와 상호작용하여 조절효과를 가지는지를 실 증한 결과이다. 분석결과 모형 4의 적합성은 통계적으로 유의하였으며(F=61.849, p<.001), 이는 모형 3에 비해 0.7%(p<.05) 증가하여 종속변수와 독립변수의 관계에서 조절변수가 조 절효과가 있음을 보여준다.

이러한 결과는 독립변수가 높아지면 종속변수는 증가하고, 조절변수가 증가할수록 종속변 수는 감소하는 것을 의미한다. 또한 독립변수와 조절변수가 상호작용함으로써 종속변수의 영향력이 높아짐을 보여준다.

## 3. 회귀분석을 활용한 매개효과 분석

조절효과분석을 해보았으니, 이제 매개효과분석도 할 줄 알아야겠지요? 종속변수와 독립변수 사이에서 매개를 하는 변수가 매개변수입니다. 매개효과는 매개변수가 통계 적으로 유의미하게 매개를 하는지 하지 않는지를 알아보는 분석입니다.

회귀분석을 활용한 매개효과분석은 어렵지는 않지만, 약간 복잡합니다. 여러번 회귀 분석을 돌려야 하기 때문이죠.

회귀분석을 활용한 매개효과분석 순서

- 1. 독립변인 → 매개변인
- 2. 독립변인 → 종속변인
- 3. 독립변인 + 매개변인 → 종속변인

위와 같은 순서로 회귀분석을 돌리시면 됩니다.

🚺 매개효과생	🔚 매개효과샘플.sav [데이터집합3] - IBM SPSS Statistics Data Editor									
파일(F) 편집	입(E) 보기(V) 데이	터(D) 변환(T)	분석(A) 다이렉트 마케팅	(M)	그래프	( <u>G</u> ) 유틸리티	티(U) 창(W) 도			
		5	보고서(P) 기술통계량(E)	×	1	*5				
			Ŧ							
	ID	독립변인	평균 비교(M)	*	변인	변수	변수			
1	1	3.00	일바셔현모현(G)		3.50					
2	2	3.78	이바킹 세칭 모형/7\		2.75					
3	3	3.33	글린와 신영 또영(Z)		2.50					
4	4	3.11	온탑 모형( <u>X</u> )	•	3.00					
5	5	3.33	상관분석( <u>C</u> )	•	2.75					
6	6	3.00	회귀분석( <u>R</u> )	•	📃 자동	동 선형 모형화.				
7	7	3.00	로그선형분석( <u>O</u> )		- 전형	열(L)				
8	8	3.11	신경망( <u>W</u> )	•	· · · · · · · · · · · · · · · · · · ·	- 1초전(C)				
9	9	3.22	분류분석(Y)	•						
10	10	3 67	차위 가소(D)			루죄소세곱(S)				

분석(A) → 회귀분석(R) → 선형(L)을 클릭하여서 "선형 회귀분석" 창을 엽니다. 이 작업은 계속합니다.

🕼 선형 회귀분석	×
종속변수(D):         ● 매개변인         ● 특립변우(I):         ● 특립변우(I):         ● 독립변인         ····································	통계량(S) 도표(T) 저장(S) 옵션(Q) 붓스트랩(B)

첫 번째로, 종속변수에 매개변인을 넣고 독립변수에 독립변인을 넣고 "확인"을 클릭하면 아웃풋이 나옵니다.

분산분석^a

모형		제곱합	자유도	평균 제곱	F	유의확를
1	회귀 모형	15.077	1	15.077	145.231	.000 ^b
	잔차	20.763	200	.104		
	합계	35.840	201			

a. 종속변수: 매개변인

b. 예측값:(상수), 독립변인

계수ª

	비표준화 계수		표준화 계수		
모형	В	표준오차	베타	t	유의확를
1 (상수)	1.622	.165		9.802	.000
독립변인	.546	.045	.649	12.051	.000

a. 종속번수: 매개번인

분산분석결과 모형이 유의하였으며, 독립변인이 매개변인에 유의확률 .000으로 통계적 으로 유의미한 영향을 미치고 있습니다.



두 번째로, 종속변수에 종속변인을 넣고 독립변수에 독립변인을 넣고 "확인"을 클릭하면 아웃풋이 나옵니다.

분산분석^a

모형		제곱함	자유도	평균 제곱	F	유의확를
1	회귀 모형	10.056	1	10.056	22.056	.000 ^b
	잔차	91.186	200	.456		
	합계	101.243	201			

a. 종속변수: 종속변인

b. 예측값:(상수), 독립변인

ואר	лa
-11	-

		비표준화 계수		표준화 계수		
모형		В	표준오차	베타	t	유의확를
1	(상수)	1.483	.347		4.278	.000
	독립변인	.446	.095	.315	4.696	.000
a. <del>a</del>	통속변수:종속	변인				

분산분석결과 모형이 유의하였으며, 독립변인이 종속변인에 유의확률 .000으로 통계적 으로 유의미한 영향을 미치고 있습니다.

🕼 선형 회귀분석	×
●● ID ●● 독립변인 ●● 매개변인	종속변수(D):
	방법(M): 입력 < 선택변수(C): 규칙(U)
	케이스 설명( <u>C</u> ): WLS 가중값( <u>H</u> ):
확인	붙여넣기(P) 재설정(R) 취소 도움말

세 번째로, 종속변수에 종속변인을 넣고 독립변수에 독립변인와 매개변인을 넣고 "확인" 을 클릭하면 아웃풋이 나옵니다.

분산분석^a

모형		제곱합	자유도	평균 제곱	F	유의확를
1	회귀 모형	12.431	2	6.216	13.928	.000 ^b
	잔차	88.811	199	.446		
	합계	101.243	201			

a. 종속변수: 종속변인

b. 예측값:(상수), 매개변인, 독립변인

	비표준화 계수		표준화 계수		
	В	표준오차	포타	t	유의확를
수)	.935	.417		2.239	.026
] 변인	.261	.123	.185	2.117	.036
개변인	.338	.147	.201	2.307	.022
a. 종속변수: 종속변인					
	수) 웹번인 배번인 수:종속	비표준 : B 수) .935 입변인 .261 배변인 .338 수: 중속변인	비표준화 계수 B 표준오차 수) .935 .417 입법인 .261 .123 배번인 .338 .147 수: 종속번인	비표준화계수 표준화계수 B 표준오차 베타 수)	비표준화계수         표준화계수           B         표준오차         베타         t           수)         .935         .417         2.239           실번인         .261         .123         .185         2.117           배번인         .338         .147         .201         2.307

계수ª

분산분석결과 모형이 유의하였으며, 독립변인과 매개변인이 종속변인에 통계적으로 유의 미한 영향을 미치고 있습니다.

이를 표로 만들어서 해석을 하면 다음과 같습니다.

### 예제) 회귀분석을 활용한 매개효과과분석

종속변인와 독립변인에서 매개변수가 매개효과가 있는지 알아보기 위해서 Baron과 Kenny 가 제시한 단계에 따라 위계적 다중회귀분석을 실시하였다. 매개효과를 검증하기 위해서는 다음의 3가지 조건을 만족시켜야 한다.

첫째, 독립변인이 매개변인에 영향을 주어야 하며(경로A),

둘째, 독립변인은 종속변인에 유의한 영향을 주어야하고(경로B),

셋째, 독립변인과 매개변인이 동시에 종속변인에 유의미한 영향을 주는지 검증했을 때 2단 계에서 유의미한 관계였던 독립변수와 종속변인의 관계가 약해지거나(부분매개) 유의하지 않아야 한다(완전매개).

이상에서 제시한 단계에 따라 매개변인에 대한 타당성을 분석한 결과는 다음과 같다.

단계	도립	조소	비표준화 계수		표준화 계수	t	유의 확률	
			β	표준오자	베타			
1단계	독립변인	매개변인	0.546	0.045	0.649	12.051	0.000	
(경로A)		R ² =.421, F=145.231***						
2단계	독립변인	종속변인	0.446	0.095	0.315	4.696	0.000	
(경로B)		R ² =.099, F=22.056***						
3단계	독립변인	종속변인	0.261	0.123	0.185	2.117	0.036	
(경로C)	매개변인		0.338	0.147	0.201	2.307	0.022	
			R ² =.123, F=13.928***					

[표] 종속변인와 독립변인에서 매개변인의 매개효과

***p<.001

1단계에서 독립변인이 매개변인을 유의미하게 예측하는지 검증한 결과 독립변인이 매개변 인에 유의미한 영향을 미치는 것으로 나타나(β=.546, p<.001), 첫째 조건을 충족시켰다. 2단계에서 독립변인이 종속변인을 유의미하게 예측하는지 검증한 결과 독립변인이 종속변 인에 유의미한 영향을 미치는 것으로 나타나(β=.446, p<.001), 두 번째 조건을 충족시켰다. 3단계에서 독립변인과 매개변인을 동시에 투입하여 종속변인에 미치는 영향력을 검증하였 다. 독립변인과 매개변인 모두 종속변인에 유의미한 영향을 미쳤다. 독립변인이 종속변인에 미치는 영향력은 2단계에 비하여 감소하였다(2단계: β=.446, 3단계: β=0.261). 이는 매개 변인이 독립변인과 종속변인 사이에서 부분 매개효과를 한다고 볼 수 있다.

마지막으로 독립변인이 매개변인을 통해 종속변인에 미치는 간접효과에 대한 유의도를 검 증하기 위해 Sobel-test를 실시한 결과 매개변인의 간접효과(β=.295)는 유의함(Sobel's T: Z=4.446, p<.05)을 나타냈다. 따라서 매개변인은 독립변인이 종속변인에 미치는 영향을 부 분 매개하는 것으로 볼 수 있으며, 이와 관련된 부분 매개효과 모형은 다음과 같다.


속변인과 독리변인의 관계에서 매개변인의 매개효과 모영 ( ): 매개변인의 간접효과

http://www.danielsoper.com/statcalc/calculator.aspx?id=31

위 사이트에 방문하면 Sobel-test를 할 수 있습니다.

회귀분석을 활용한 매개효과분석 명령어(Syntax) REGRESSION /MISSING LISTWISE /STATISTICS COEFF OUTS R ANOVA /CRITERIA=PIN(.05) POUT(.10) /NOORIGIN /DEPENDENT 매개변인 /METHOD=ENTER 독립변인. REGRESSION /MISSING LISTWISE /STATISTICS COEFF OUTS R ANOVA /CRITERIA=PIN(.05) POUT(.10) /NOORIGIN /DEPENDENT 종속변인 /METHOD=ENTER 독립변인. REGRESSION /MISSING LISTWISE /STATISTICS COEFF OUTS R ANOVA /CRITERIA=PIN(.05) POUT(.10) /NOORIGIN /DEPENDENT 종속변인 /METHOD=ENTER 독립변인 매개변인.

# 여기서

/DEPENDENT 뒤에는 종속변인 이름으로 바꾸고 /METHOD=ENTER 뒤에는 독립변인 이름을 넣으면 됩니다. (.) 명령어 맨뒤에 점은 꼭 넣어주세요.

## 4. 이항로지스틱 회귀분석

회귀분석의 종속변수가 연속형이었다면, 이항로지스틱 회귀분석은 종속변수가 이분형 척도 측정된 변수 예) 남성/여성, 병에 걸림/안걸림, 있다/없다 이런식으로 나눠져야 합니다. 독립은 연속형이던 비연속형이던 상관은 없습니다. 의료통계에서 대부분 사용 하고 있는 분석방법입니다.

이항로지스틱 회귀분석에서의 핵심은 EXp(B)=odd ratio로 종속변수에 독립변수가 얼마만큼 부(-)적이던, 정(+)적이던 영향을 주는지를 아는 것입니다.

	*로지스틱실	생품1.sav [데이터진한2]	- IBM SPSS Stati	stics Data Editor			
100	파일(F) 편집	임(E) 보기(V) 데이티	H(D) 변환(T)	분석(A) 다이렉트 마케팅(M)			
200	🔁 H	출금을 🔄 🖛 🤉 (보고서면) 기술통계량(E)					
				E >	-		
		선택여부	비연속형변=	평균비교( <u>M</u> ) ▶	수 변수 변수		
	1	선택함	선택요인	일반선형모형(G) ▶			
	2	선택하지 않음	선택요인	일바하셔형 모형(7) 🕨			
	3	선택하지 않음	선택요인	호한 미천 (V)			
	4	선택하지 않음	선택요인	도립 포영(V) 🕨			
	5	선택하지 않음	선택요인	상관분석( <u>C</u> ) ▶			
	6	선택하지 않음	선택요인	회귀분석( <u>R</u> ) ▶	🗾 자동 선형 모형화		
	7	선택하지 않음	선택요인	로그선형분석( <u>O</u> ) ▶	🔚 선형(L)		
	8	선택하지 않음	선택요인	신경망( <u>W</u> ) 🕨	☑ 곡선추정(C)		
	9	선택하지 않음	선택요인	분류분석(Y) ▶	이 이 비 치 시계고/아		
	10	선택하지 않음	선택요인	차원 감소( <u>D</u> ) ▶	·····································		
	11	선택하지 않음	선택요인	척도(A) 🕨	🔜 미분형 로지스틱( <u>G</u> )		
	12	선택하지 않음	선택요인	U모수 검정(N) ▶	🔜 다항 로지스틱( <u>M</u> )		
	13	선택하지 않음	선택요인		🔛 순서(D)		
	14	선택하지 않음	선택요인	새조화로(이 .	[編] 프로빗(P)		
		나 다 파	HELOOL	824 <u>2</u> (0)			

분석(A) → 회귀분석(R) → 이분형 로지스틱(G) 클릭



종속변수에 종속변수를 넣고 공변량(C)에 독립변수를 넣고 "범주형(C)"을 클릭 후 독립변수중 비연속형변수를 범주형 공변량(T)에 넣고 참조범주를 "마지막" 또는 "처음" 으로 선택해줍니다. 여기서 기본값은 "마지막"이지만 저희는 "처음"으로 바꾸도록 하겠습 니다.

	로지스틱 회귀분석: 옵션     ┌통계량 및 도표     ································	
범주형( <u>C</u> ) 저장( <u>S</u> )	<ul> <li>✔ 분류도표(C)</li> <li>✔ Hosmer-Lemeshow 적합도(H)</li> </ul>	<ul> <li>추정값들의 상관계수(R)</li> <li>한복계산정보(I)</li> </ul>
옵션(0) 붓스트랩(T)	<ul> <li>□ 케이스별 잔차목록(W):</li> <li>● 밖에 나타나는 이상값(○) 2 표</li> <li>● 전체 케이스(A)</li> <li>- 표시</li> <li>● 강 단계마다(E) ● 마지막 단계에 서(1)</li> </ul>	✔ exp(B)에 대한 신뢰구간(X): 95 %의 케이스 추출 중편차
	단계선택에 대한 확률 진입(E): 0.05 제거(A): 0.10	분류 분리점(U): 0.5 최대반복계산수(M): 20
말	<ul> <li>■ 복잡한 분석 또는 큰 데이터 집합을 위해 □</li> <li>▼ 모형에 상수 포함(!)</li> </ul>	제국 취소 도움말

다음에는 옵션(O)를 클릭하여 "분류도표(C)", "Hosmer-Lemeshow 적합도(H)", "exp(B)에 대한 신뢰구간 95%"를 클릭하고 "계속"을 클릭하여 아웃풋을 출력합니다.

케이스 처리 요약

가중되지 않은 케이스 ^a	Ν	퍼센트
선택케이스 분석에포함	299	99.7
결측 케이스	1	.3
합계	300	100.0
비선택 케이스	0	.0
합계	300	100.0

a. 가중값을 사용하는 경우에는 전체 케이스 수의 분류표를 참조하십시오.

종속변수 코딩

원래 값	내부 값
선택하지 않음	0
선택함	1

범주형 변수 코딩

			파러미터 코딩
		빈도	(1)
비연속형변수	선택요인 없음	117	.000
	선택요인 있음	182	1.000

종속변수의 코딩은 선택함이 1번으로, 범주형 독립변수 코딩은 선택요인 있음이 1번으로 선택되었습니다.

# 블록 0: 시작 블록

				예측	
			선택여	₽	
	감사됨		선택하지 않음	선택함	분류정확 %
0 단계	선택여부	선택하지 않음	0	90	.0
		선택함	0	209	100.0
	전체 퍼센트	Ē			69.9

a. 모형에 상수항이 있습니다.

b. 절단값은 .500입니다.

선택하지 않음 집단 90명과 선택함 집단 209명은 모두 선택함 집단에 분류되었고 분류 정확도는 69.9%였습니다.

#### 방정식에 포함된 변수

	В	S.E,	Wals	자유도	유의확를	Exp(B)
0단계 상수항	.843	.126	44.656	1	.000	2.322

방정식에 포함되지 않은 변수

			점수	자유도	유의확를
0 단계	변수	비연속형변수(1)	178.952	1	.000
		연속형변수	185.517	1	.000
	전체 통계량			2	.000

블록0(시작단계)에서는 독립변수를 제외한 상수항만으로 구성된 식이 나타납니다. 모형에 포함되지 않은 독립변수의 유의확률도 확인할 수 있습니다.

# 블록 **1:** 방법 **=** 진입

모형 계수 전체 테스트

		카이제곱	자유도	유의확를
1 단계	단계	270.299	2	.000
	볼록	270.299	2	.000
	모형	270.299	2	.000

모형 계수 전체 데스트표는 독립변수들이 포함되었을 때의 유용성을 보여주는 표로 여 기서는 카이제곱값이 270.299, 유의확률이 .000으로 유의수준 .05보다 작기 때문에 두 개의 독립변수가 선택유무를 판단하는데 유용하다고 할 수 있습니다.

모형 요약

단계	-2 Log 우도	Cox와 Snell의 R-제곱	Nagelkerke R-제곱
1	95.505 ^a	.595	.843

a. 모수 추정값이 .001보다 작게 변경되어 계산반복수 7에서 추정을 종료하였습니다.

= Hosmer 와 Lemeshow 검정 =

단계	카이제곱	자유도	유의확를
1	5.749	6	.452

모형요약에 Nagelkerke R제곱값은 회귀분석의 R제곱값과 같이 모형의 설명력을 나타 냅니다. 여기서는 84.3%의 모형 설명력을 가지고 있네요.

Hosmer와 Lemeshow 검정은 모형의 적합도 검정으로 유의확률이 .05보다 크면 적합 도가 좋다고 해석합니다. 여기서는 .452이니깐 모형이 적합하다고 해석합니다.

분류표ª

			선택여!		
	감사됨		선택하지 않음	선택함	분류정확 %
1 단계	선택여부	선택하지 않음	87	3	96.7
		선택함	13	196	93.8
	전체 퍼센트	Ē			94.6

a. 절단값은 .500입니다.

분류표에서는 정확히 예측한 숫자가 얼마만큼 되는 알려주는 표로 전체 분류정확도는 94.6%로 나타났습니다.

방정식에 포함된 변수

								EXP(B)에 대한	95% 신뢰구간
		В	S.E,	Wals	자유도	유의확를	Exp(B)	하한	상한
1 단계 ^a	비연속형변수(1)	3.028	.694	19.054	1	.000	20.649	5.303	80.408
	연속형변수	2.505	.437	32.870	1	.000	12.249	5.201	28.846
	상수항	-7.592	1.208	39.483	1	.000	.001		

a. 변수가 1: 단계에 진입했습니다 비연속형변수, 연속형변수. 비연속형변수, 연속형변수.

가장 중요한 방정식표입니다.

B의 부호가 +이면 변수값이 클수록 내부값이 1인 여기서는 "선택함" 포함될 확률이 크 며, 부호가 -이면 변수값이 클수록 내부값이 0인 "선택하지 않음"에 포함될 확률이 커집 니다.

이 표를 해석하려면, Exp(B)값과 유의확률을 살펴봐야합니다. 모든 독립변수의 유의확 률은 .000으로 모두 통계적으로 유의미했으며, 비연속형변수(1)의 Exp(B)는 20.649로 여 기서 비연속형변수(1)은 "선택요인 있음" 항목으로 선택요인 있음이 있을 때 선택함에 속활 확률이 20.649배가 된다는 뜻이고, 연속형변수 Exp(B)는 12.249로 연속형변수가 1 증가할 때 선택함에 포함될 확률이 12.249배 커진다고 해석합니다.

Exp(B)의 신뢰구간 하한과 상한사이에 1이 포함되지 않아야 통계적으로 유의미한 독립 변수가 됩니다. 5. 공분산분석

공분산분석은 영향을 줄 수 있는 공변인을 통제하여 독립변수가 종속변수에 순수한 영향을 주는지를 알아보는 분석방법입니다.

보통 두 집단의 사전-사후 점수의 차이를 알아볼 때 사전 점수를 공변인으로 통제하 고 사후점수를 종속변수로 선택하여 분석합니다.

공분산분석에 있어서 사전점수가 동질하다면 굳이 공분산분석을 실시하지 않아도 됩니다. 또한 집단과 사전검사점수의 상호작용효과가 통계적으로 유의미할 때도 공분산 분석을 하지 않아도 됩니다.

*공분	산샘플.	sav [데이터집합	9] - IBM SPSS Stat	istic	s Data Ec	litor	-						
파일( <u>F</u> )	편집	(E) 보기(V)	데이터(D) 변환(	<u>(</u> )	분석( <u>A</u> )	다이렉트디	과케팅(M	<u>/)</u>	그래프( <u>G</u> )	유틸리	EI(U)	창( <u>W</u> )	
					보고 기술 표	1서(P) š통계량( <u>E</u> )	۲ ۲		1			4	
		집단	사전_몸무게	٨ł	평균	문 비교(M)		ſ	지다벽	평규부석	(M)		
1		비교집단	61.4		일반	반선형모형(G	s) Þ		Полн	- 775	( <u>m</u> )		
2		비교집단	52.1		인터	바하 세형 모형	di(7) ⊾		1 일표본	1 1 2 8 8	)	_	
3		비교집단	59.7			:피근ᇢㅗ;	-(2)		🛃 독립표본 <u>T</u> 검정(T)				
4		비교집단	60.2		~ ~ ~	1 포영( <u>V</u> )	P		🔛 대응표	본 T 검정(	( <u>P</u> )		
5		비교집단	68.4		상품	*문역( <u>C</u> )			[ 일원배	치 분산분	석(0)		
6		비교집단	58.2		회7	분석( <u>R</u> )	1	ł					
7		비교집단	78.3		로 2	1선형분석( <u>0</u>							
(mm)													

사전점수 동질성을 알아보기 위해

분석(A) → 평균비교(M) → 독립표본 T검정(T) 클릭



사전_몸무게 검정변수(T)에 넣고 집단을 정의해줍니다.

집단통계량

	집단	N	평균	표준편차	평균의 표준오차
사전_몸무게	비교집단	34	64.429	9.2857	1.5925
	실험집단	30	60.053	7.7789	1.4202

Levene의 동분산 검정				평균의 동일성에 대한 t-검정						
						유의확를		차이의	차이의 959	6신뢰구간
		F	유의확를	t	자유도	(양쪽)	평균차	표준오차	하한	상한
사전_몸무게	등분산이 가정됨	2.868	.095	2.028	62	.047	4.3761	2.1577	.0630	8.6892
	동분산이 가정되지 않음			2.051	61.848	.045	4.3761	2.1338	.1105	8.6417

Levens의 등분산이 가정되어 사전_몸무게는 유의확률 .047로 통계적으로 유의미한 차이가 나타났습니다. 그럼 한가지 요건이 충족되었습니다.

ta *공분산샘플.sav [데이터집합9] - IBM SPSS Statistics Data Editor									
파일(F) 편	집(E) 보기(V)	데이터(D) 변환(	[)	분석( <u>A</u> )	다이렉트 마케팅	∃( <u>M</u> )	그래프(G)	유틸리	EI(U
				보고 기술 표	1서(P) 호통계량(E)	•			
	집단	사전_몸무게	사	표평균	원 비교(M)	*	변수	변수	
1	비교집단	61.4		일반	·선형모형(G)			สม	
2	비교집단	52.1		일빈	· 한화 선형 모형(Z)			<u>(</u> )	
3	비교집단	59.7		Żź	년 모현(X)			( <u>M</u> )	
4	비교집단	60.2					🏭 반복측	정( <u>R</u> )	
5	비교집단	68.4		85	r=⊣( <u></u> )	P	분산성	분(V)	
6	비교집단	58.2		회7	분석( <u>R</u> )	1			
7	비교집단	78.3		로그	1선형분석( <u>O</u> )	1			
8	비교집단	48.1		신경	명망(W)	•			

이제 집단과 사전검사점수의 상호작용효과가 있는지 살펴보겠습니다. 분석(A) → 일반선형모형(G) → 일변량(U) 클릭



종속변수(D)에는 "사후_몸무게"를 모수요인(F)에는 "집단"을 공변량(C)에는 "사전_몸무 게"를 넣고 모형(M)을 클릭합니다. 사용자정의(C)를 클릭하고 집단, 사전_몸무게를 클릭 해서 오른쪽으로 옮기고 집단과 사전_몸무게를 동시에 클릭해서(Shift사용) 오른쪽으로 옮깁니다. 그리고 "계속"을 클릭

개체-간 효과 검정

종속 변수: 사후_몸무게

	제 Ⅲ 유형 피고착	лос	편크 피고	F	0 0 <b>0 0</b>
소스	제곱암	자규도	평균 제곱	Г	<u> </u>
수정 모형	4840.345 ^a	3	1613.448	1446.716	.000
절편	.099	1	.099	.089	.767
집단	.056	1	.056	.050	.824
사전_몸무게	4211.448	1	4211.448	3776.240	.000
집단*사전_몸무게	.003	1	.003	.003	.959
오차	66.915	60	1.115		
합계	248955.610	64			
수정 합계	4907.260	63			

a. R 제곱 = .986 (수정된 R 제곱 = .986)

집단과 사전_몸무게의 유의확률이 .959로 통계적으로 유의미하지 않았기 때문에 상호작 용 효과가 없습니다. 이는 공분산분석을 해도 된다는 이야기입니다.



이번에는 모형(M)을 클릭하여 완전요인모형(A)를 클릭합니다.



"옵션(O)"를 클릭해서 "효과크기 추정값(E)", "모수추정값(T)", "동질성 검정(H)" 체크해 주고 아웃풋을 돌립니다.

#### 오차 분산의 동일성에 대한 Levene의 검정^a

종속 변수: 사후_몸무게

F	df1	df2	유의확를					
.697	1	62	.407					
여러 집단에서 종속변수의 오차 분산이 등일한								

영가설을 검정합니다.

a. Design: 절편 + 사전_몸무게 + 집단

개체-간 효과 검정

종속 변수: 사후_몸무게

	제 III 유형					
소스	제곱함	자유도	평균 제곱	F	유의확를	부분 에타 제곱
수정 모형	4840.342 ^a	2	2420.171	2206.143	.000	.986
절편	.096	1	.096	.087	.768	.001
사전_몸무게	4463.996	1	4463.996	4069.222	.000	.985
집단	4.498	1	4.498	4.101	.047	.063
오차	66.918	61	1.097			
합계	248955.610	64				
수정 합계	4907.260	63				

a. R 제곱 = .986 (수정된 R 제곱 = .986)

Levens의 등분산이 유의수준 .407로 가정되었습니다.

개체-간 효과 검정을 살펴보면, 사전_몸무게는 유의확률 .000으로 몸무게 차이는 사전_ 몸무게에 따라 차이가 날수도 있다고 나왔습니다.

집단은 유의확률 .047로 집단에 따라 몸무게 차이가 나타났습니다.

#### 모수 추정값

종속 변수: 사후_몸무게

						95% 신	뢰구간	
	모수	В	표준오차	t	유의확를	하한값	상한값	부분 에타 제곱
۰	절편	.013	.947	.013	.989	-1.881	1.906	.000
	사전_몸무게	.985	.015	63.790	.000	.954	1.016	.985
	[집 단=1]	.549	.271	2.025	.047	.007	1.090	.063
	[집 단=2]	0ª						

a. 이 모수는 중복되었으므로 0으로 설정됩니다.

모수 추정값을 살펴보면, 비교집단과 실험집단의 사후검사의 평균은 사전_몸무게가 통제 된 후에 .549의 차이가 나타났으며, 이는 유의확률이 .047로 유의수준 .05보다 작아 통 계적으로 유의미한다고 해석합니다. 예제) 공분산분석

헬스클럽에서 운동을 실시 한 실험집단과 헬스클럽을 다니지 않은 대조집단의 몸무 게 변화를 비교하기 위해 사전 몸무게를 공변인으로 하고 사후 몸무게를 종속변인으 로 하여 공분산 분석을 실시하였다.

실험집단과 비교집단의 몸무게 사전-사후 평균과 표준편차는 다음표에 제시하였
 다.

실험집단의 몸무게는 사전 평균 60.05kg이었으며, 사후 59.17kg로 약 0.88kg 줄었으며, 비교집단의 몸무게는 사전 평균 64.43kg이었으며, 사후 64.03kg로 약 0.40kg 줄은 것으로 나타났다.

[표] 실험실단과 비교집단의 몸무게 사전-사후 평균과 표준편차

୦ ୦]	사전사	실험집단	비교집단
<u>ــــــــــــــــــــــــــــــــــــ</u>	후	평균 ± 표준편차	평균 ± 표준편차
모 ㅁ -))	사전	$60.05 \pm 7.78$	$64.43 \pm 9.29$
고구계	사후	$59.17 \pm 7.73$	$64.03 \pm 9.21$

두 집단의 헬스클럽을 다니기 전과 후의 몸무게 차이가 있는지 통계적으로 살펴보기 위해 공분산분석을 한 결과는 다음과 같다.

오차분산의 동일성에 대해 알아보기 위해 Leven의 검정을 실시한 결과는 다음과 같다.

[표] Leven의 검정

F	df1	df2	유의확률
.697	1	62	.407

Leven 검정결과 유의확률이 .05이상으로 등분산이 가정되었다.

[표] 몸무게의 공분산 분석 결과

변량원	자승합	자유도	평균자승	F	р
사전검사	4463.996	1	4463.996	4069.222	.000
집단	4.498	1	4.498	4.101	.047
오차	66.918	61	1.097		

위 표에서 보는 바와 같이 몸무게는 사후 검사에서 실험집단과 비교집단은 통계적으 로 유미의미한 차이가 나타났다(p<.05). 헬스클럽을 다녔던 실험집단의 몸무게가 비교 집단의 몸무게보다 좀 더 많이 줄어든 것을 알 수 있었다.



## 6. 반복측정 분산분석

의료통계에서 많이 사용되는 반복측정 분산분석입니다.

실험집단과 비교집단의 실험처치 후 여러번 측정된 값들의 차이가 있는지를 알아보 는데 많이 사용합니다.

🚺 반복측정심	🚂 반복측정샘플데이터.sav [데이터집합1] - IBM SPSS Statistics Data Editor										
파일(F) 편	집(E) 보기(V)	데이터(D) 법	변환(T)	분석( <u>A</u> )	다이렉트 마케팅	(M)	그래프( <u>G</u> )	유틸리	IEI(U		
					1서(P) 출통계량(E)	•	H 🖁		4		
		-	Ŧ		•	2					
	ID	집단	l	평균	원비교(M)	•	체중_2회	ĺ į	j		
1	1	대조군		일반	·선형모형(G)	F.	111 안벼루(JI	1			
2	2	대조군		일반화 선형 모형(Z) ▶ 혼합 모형(X) ▶ 상관분석(C) ▶		의바하셔형 모형(7) 🕨 📼 드					
3	3	대조군				₩ 나면당( <u>M</u>	)				
4	4	대조군							1 반복측정	( <u>R</u> )	
5	5	대조군				P	분산성분	(⊻)			
6	6	대조군		회7	· 분석( <u>R</u> )	1		<u>69.2</u>			
7	7	대조군		로그	1선형분석( <u>O</u> )	1		68.2			

분석(A) → 일반선형모형(G) → 반복측정(R) 클릭



그럼 반복측정 요인 정의창이 먼저 뜹니다. 여기서 "개체-내 요인이름(W)"에는 보통 "시 기", "측정차시"와 같이 시간과 관련된 이름을 넣어주고 "추가(A)"를 클릭하고 "정의"를 클릭합니다.



"개체-내 변수(W)"에는 여러번 측정된 데이터값을 "개체-간 요인(B)"에는 집단을 넣어 주고 도표(T)를 클릭하고 "집단"을 선구분 변수(S)에 "시기"를 수평축 변수(H)에 넣고 추가를 누르고 "계속"을 클릭하고 "확인"을 클릭하여 아웃풋을 생성시킵니다.

#### Mauchly의 구형성 검정^a

축도: MEASURE_1							
						옙실런 ^b	
					Greenhouse-		
개체-내 효과	Mauchly의 W	근사 카이제곱	자유도	유의확를	Geisser	Huynh-Feldt	하한값
측정시기	.760	7.335	5	.197	.865	.995	.333

정규화된 변형 종속변수의 오차 공분산행렬이 단위행렬에 비례하는 영가설을 검정합니다.

a. Design: 절편 + 집단

개체-내 계획: 측정시기

b. 유의성 평균검정의 자유도를 조절할 때 사용할 수 있습니다. 수정된 검정은 개체내 효과검정 표에 나타납니다.

우선 Mauchly의 구형성 검정결과 유의확률이 .197로 유의수준 .05보다 커 구형성의 가 정을 만족합니다.

* Levens와 같이 유의수준이 .05보다 클 때 등분산성이 성립한다와 같다.

Mauchly의 구형성 검정결과 만약 유의확률이 .05 미만일 경우도

Greenhouse-Geisser, Huynh-Feldt 두 엡실런값이 0.7이상이면, 엡실런 수정법을 사용 하여 검정하면 됩니다. 구형성이 가정되었다면, 개체-내 효과 검정에서 구형성 가정 값을 보고 해석합니다.

개체-내 효과 검정

측도:	MEASURE 1
	_

소스		제 Ⅲ 유형 제곱합	자유도	평균 제곱	F	유의확를
측정시기	구형성 가정	26.940	3	8.980	18.123	.000
	Greenhouse-Geisser	26.940	2.595	10.382	18.123	.000
	Huynh-Feldt	26.940	2.985	9.024	18.123	.000
	하한값	26.940	1.000	26.940	18.123	.000
측정시기 * 집단	구형성 가정	27.885	3	9.295	18.758	.000
	Greenhouse-Geisser	27.885	2.595	10.746	18.758	.000
	Huynh-Feldt	27.885	2.985	9.341	18.758	.000
	하한값	27.885	1.000	27.885	18.758	.000
오차(측정시기)	구형성 가정	41.623	84	.496		
	Greenhouse-Geisser	41.623	72.654	.573		
	Huynh-Feldt	41.623	83.590	.498		
	하한값	41.623	28.000	1.487		

여기서는 구성형 가정이 되었으니, 구형성 가정값을 살펴보면, 시기와 시기*집단 모두 유의확률이 .000으로 유의수준 .05보다 작아 통계적으로 유의미한 차이가 있었습니다. 이를 해석하면 시기의 경우 측정시기에 따라 체중의 차이가 있다는 것을 알 수 있으며, 시기*집단의 상호작용효과가 존재한다. 즉, 두 집단에 대한 평균체중이 측정기간에 따라 달라진다는 것을 알 수 있습니다.

다	번	랑	검	정리
---	---	---	---	----

효과		값	F	가설 자유도	오차 자유도	유의확를
측정시기	Pillai의 트레이스	.618	13.994 ^b	3.000	26.000	.000
	Wilks의 람다	.382	13.994 ^b	3.000	26.000	.000
	Hotelling의 트레이스	1.615	13.994 ⁶	3.000	26.000	.000
	Roy의 최대근	1.615	13.994 ^b	3.000	26.000	.000
측정시기 * 집단	Pillai의 트레이스	.686	18.918 ^b	3.000	26.000	.000
	Wilks의 람다	.314	18.918 ^b	3.000	26.000	.000
	Hotelling의 트레이스	2.183	18.918 ⁶	3.000	26.000	.000
	Roy의 최대근	2.183	18.918 ^b	3.000	26.000	.000

a. Design: 절편 + 집단

개체-내 계획: 측정시기

b. 정확한 통계량

구형성이 가정되지 않고, 엡실런 값이 0.7 미만일 경우 "다변량 검정"을 가지고 검정하 게 됩니다.

Pillai의 트레이스 : 표본크기가 작거나 공분산이 동일하지 않고 집단크기 차이 있을 때 Wiks의 람다 : 표본크기가 충분하고 가정도 어느정도 충족하고 집단크기가 유사할 때 Roy의 최대근 : 가장 보수적, 모든 가정을 가장 엄격하게 충족시킬 때

여기서는 집단의 크기가 유사하기 때문에 Wilks의 람다 값을 가지고 해석합니다.

### 예제) 반복측정분산분석

본 연구는 두 집단(실험집단, 비교집단) 체지방량이 프로그램을 실시하기 전을 기준 으로 4일 동안 4번의 측정으로 변화율에 어떠한 변화가 있었는지에 대해 알아보고자 했다. 다음은 4일 동안 체지방량의 변화에 대한 결과이다.

실험집단의 경우 체지방량은 1차 측정 평균 23.98에서 실험 후 4일(4차 측정)에서 평균 21.31로 약 2.60점만큼 체지방량이 감소하였으며, 비교집단의 경우 1차 측정 평 균 24.99에서 실험 후 4일(4차 측정) 평균 25.11로 약 0.12점만큼 체지방량이 증가 하였다. 다음 그림을 살펴보면 실험집단이 비교집단보다 체지방량이 더 감소하는 것으 로 나타났다.

[표] 집단에 따른 체중 반복측정 평균 및 표준편차

추저 키시	실험	집단	비교집단			
특성 자시	Mean	SD	Mean	SD		
체지방량 1차 측정	23.98	4.52	24.99	5.36		
체지방량 2차 측정	23.13	4.49	24.92	5.79		
체지방량 3차 측정	22.20	4.06	24.64	5.12		
체지방량 4차 측정	21.31	4.14	25.11	5.62		



[그림] 체중 반복측정 평균 그래프

집단간 차이를 알아보기 위해 구체적으로 본 실험의 Maychly의 구형성 검정을 실시 한 결과는 다음과 같다.

[표] Maychly의 구형성 검정

개체-		지유	윤이	1실련			
내 효과	Mauchly's W	근사카이제곱	도	비ㅋ 확률	Greenhouse- Geisser	Huynh-Feldt	Lower-bound
측정 차시	.760	7.335	5	.197	.865	.995	.333

유의확률이 .197로 유의수준 .05보다 높아 통계적으로 구성형 조건이 만족하였다. 이 경우 구형성 가정값을 사용하여 집단 간 차이를 검정하기로 한다.

[표 87] 개체-내 효과 검정

소스	엡실런	제 III 유형 제곱합	자유도	평균제곱	F	유의 확율
	구형성 가정	26.940	3	8.980	18.123	.000
측정	Greenhouse-Geisser	26.940	2.595	10.382	18.123	.000
차시	Huynh-Feldt	26.940	2.985	9.024	18.123	.000
	Lower-bound	26.940	1.000	26.940	18.123	.000
측정	구형성 가정	27.885	3	9.295	18.758	.000
차시	Greenhouse-Geisser	27.885	2.595	10.746	18.758	.000
*	Huynh-Feldt	27.885	2.985	9.341	18.758	.000
집단	Lower-bound	27.885	1.000	27.885	18.758	.000

측정 차시에 따라 체지방량 구형성 가정의 유의확률이 .000으로 측정 차시에 따라 체지방량이 통계적으로 유의하게 변화하였으며(p<.001), 상호작용효과(측정 차시 * 집 단) 역시 통계적으로 매우 유의하기 때문에 집단과 측정 차시 사이에 상호작용효과가 존재하였다. 즉, 두 집단에 대한 평균 체지방량은 측정 차시에 따라 달라졌다 (p<.001). 이를 정리하면 실험집단이 비교집단보다 체지방량이 더 많이 감소하였다는 것을 알 수 있었다. 7. 정규성 검정

앞에서 다루었던 모든 통계는 수집된 자료가 정규분포를 따른다는 가정에서 분석을 실시하였습니다.

중심극한정리에 따르면 표본의 크기가 30 이상이거나 모집단이 이론적으로 정규분포 를 따른다면 표본평균은 정규분포를 따른다고 가정합니다. 하지만 표본의 크리가 30 이 되지 않거나 표본의 크리가 크더라도 특정 범위에 몰려 있는 경우, 마지막으로 표 본 크기가 30은 넘더라도 정규성 검정을 실시하였냐고 질문이 들어 온다면 이에 대해 서 자신있게 말할 수 있어야 합니다.

🔚 *샘플데이태	터.sav [데이터집합1] -					1 2122						
파일(F) 편집	집(E) 보기(V) 데이	) 터(D) 변환(T)	분석( <u>A</u> )	다이렉트 마케	팅( <u>M</u> )	그래프( <u>G</u> )	유틸리티( <u>U</u> )	창( <u>W</u> )	도움말(번)			
			보고	1서( <u>P)</u>	*	H				A		
			기술	울통계량(E)	*	172 빈도분석(F)				14		
			Ŧ		•	🔚 기술통계	(D)					
	NO	성별	평균	7 비교(M)		A CHOIEN	탈색(E)		연령더미1			
		чы М	일번	일반선형모형( <u>G</u> )		교차부식		н		1.00		
5	5		일빈	반화 선형 모형(Z	) 🕨			H		1 00		
6	6	며성	혼합	혼합 모형( <u>X</u> ) (		혼합 모형(X) 🔹 🔤 🔤				H		.00
7	7	며성	상관	상관분석( <u>C</u> )	•		M <u>-</u> エエ(-)		H			
8	8	남성	회7	· 분석( <u>R</u> )	*	<u>छ</u> <u>ठ</u> -० म् न	±(Q)	H		.00		
9	9	며성	로그	1선형분석 <mark>(0</mark> )		7	0 700	H		1.00		
10	10	여성	신경	병망( <u>W</u> )		7	3 700	H		1.00		
11	11	여성	분류	루분석(Y)		7	5 <mark>700</mark>	H		1.00		
12	12	남성	차원	년 감소( <u>D</u> )		6	8 600	H		.00		

분석(A) → 기술통계량(E) → 데이터 탐색(E) 클릭



정규성을 검정할 변수를 넣고 "도포(T)"를 클릭 도표창에서 "히스토그램(H)"와 "검정과 함께 정규성도표(O)"를 클릭합니다.

정규성 검정

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	통계량	자유도	유의확를	통계량	자유도	유의확를
나이	.089	54	.200	.978	54	.438
체중	.098	54	.200	.973	54	.272

*. 이것은 참인 유의확률의 하한값입니다.

a. Lilliefors 유의확를 수정

정규성 검정 아웃풋을 보면, "Kolmogorov-Smirnov"와 "Shapiro-Wilk"값이 나옵니다.

※ Kolmogorov-Smirnov는 표본 수가 2000개 이상일 때 사용
 ※ Shapiro-Wilk는 표본 수가 2000개 미만일 때 사용

여기서는 Shapiro-Wilk값을 보도록하겠습니다.

유의확률이 0.05보다 높습니다. 이는 영가설을 기각할 수 없으므로 나이와 체중은 정규 분포를 따른다고 해석합니다.

			통계량	표준오카
11-01	평균		54.44	1 366
1.01	 평균의 95% 신립구간	하한	51.70	1.555
	0240022472	아는 산화	57.19	
	5% 전상평균	02	54.77	
			54.00	
	 분산		100.818	
			10.041	
			27	
			73	
	 		16	
			40	
	에드		200	225
	체도		300	.320
궤즈	명구	•••••	.107	.039
418	편균이 0.5% 시리그가	치하	64.15	1.435
	82433202372	사회	67.00	
	50 저사피크	35	67.03	
	- 370 월역왕간 		63.95	
	· 중취구 · · · · · · · · · · · · · · · · · · ·		62.55	
			111.132	
	_ 표준편자		10.542	
	_ 죄소값 		43	
			88	
	변원		45	
	사문위수 범위		14	
	왜도		.392	.325
	첨도		476	.639

만약 정균성 검정을 통과 하지 못한다면, 왜도와 첨도를 종합적으로 검토 합니다.

왜도와 첨도는 절대값이 2를 넘지 않으면 정규분 포로 생각합니다.

조금 느슨하게 본다면 3 을 기준으로 볼수도 있습 니다.

또한 첨도는 8까지 본다 는 논문들도 있습니다.

## 8. 비모수 통계분석

입력된 자료가 범주형 자료(명목척도, 서열척도)이거나, 입력된 데이터 크기가 30미 만일 때, 정규분포를 가정할 수 없고, 등분산성이 가정되지 않을 때는 비모수 통계를 실시해야 합니다.

비모수 통계 기법	모수 통계 기법
Mann-Whitney U 검정	독립표본 t-test
Wilcoxon 부호-서열 검정	대응표본 t-test
Kruskal-Wallis H 검정	ANOVA(분산분석)
교차분석	두 변인 모두 명목척도
Spearman 서열상관관계분석	두 변인 모두 서열척도

1) Mann-Whitney U 검정



분석(A) → 비모수검정(N) → 레거시 대화 상자(L) → 독립2-표본(2) 클릭

### - 92 -

근사 유의확률(양측)은 표본수가 30개 이상일 때 사용하고, 정확한 유의확률[2*(단측유의확률)은 표본수가 30개 미만일 때 사용하면 됩니다. 샘플은 표본수가 16개이니, 정확한 유의확률을 보면 됩니다. .038로 유의수준 .05보다 작아 통계적으로 유의미하게 차이가 있는 것을 확인할 수 있었습니다.

순위합 Ν 평균순위 집단 실험집단 87.50 8 10.94 50

12.500

48.500 -2.055

.040

.038^b

집단1과 집단2를 나누어 줍니다. 하는 방법은 독립 t-test와 같습니다. Mann-Whitney의 U(M)을 체크하고 "확인"을 클릭하면 아웃풋이 나옵니다.

비교집단	8	6.06	48
합계	16		-
건제 특	계락		
	18		
		점수	



순위

Mann-Whitney 검정

점수

유의확률)] a. 집단변수: 집단

정확한 유의확를 [2*(단측

Mann-Whitney의 U

근사 유의확률(양측)

Wilcoxon의 W

b. 등틀에 대해 수정된 사항이 없습니다.

평균순위가 높을수록 평균이 높습니다. 순위합도 마찬가지입니다.

여기서는 실험집단이 비교집단보다 평균순위가 크네요.

검정 통계량을 보시면,

2) Wilcoxon 부호-서열 검정

ta Wilcox	ion 검정.sav [데이터	집합2] - IBM S	PSS Statis	stics Data	Editor						
파일(F)	편집(E) 보기(V)	데이터(D)	변환(T)	분석( <u>A</u> )	다이렉트 마케팅	∃( <u>M</u> )	그래프( <u>G</u> )	유틸리티( <u>U</u> )	창( <u>W</u> )	도움말( <u>H</u> )	
				보고 기술 표	1서(P) š통계량(E)	*	H		4	2	( 1ର୍କ
	사전점수	사후점수	변각	ᄪ	2 HI 🗊 (M)	1	변수	변수	변각	- 변	<del>수</del>
1	85.0	66.7	1	01F	· 네표( <u>제</u> )						
2	75.0	56.7		21	2건영도영( <u>6</u> ) 바퀴 서취 미취(7)						
3	82.5	85.0	)	2t **	한화신영 포영( <u>८</u> )						
4	30.0	40.0	)	관일 사고	[모영( <u>X</u> )	1					
5	80.0	78.3	}	872 	*분석( <u>C</u> )	*					
6	87.5	75.0	)	হা ব	·[분석( <u>R</u> )	*					
7	62.5	63.3	}	로그	1선형분석 <mark>(</mark> )	*					
8	65.0	60.0		신경	뿝망( <u>W</u> )	*					
9	82.5	66.7	1	분류	루분석( <u>Y</u> )	*					
10	10.0	50.0	)	차원	년 감소( <u>D</u> )	•			_	_	
11	37.5	43.3	1	척도	E <u>(A</u> )	•					
12	35.0	41.7	1	EIR	2수 검정( <u>N</u> )	•	▲ 일표본(	D)			
13	42.5	56.7	'	예 #	ŧ(T)	•	▲ 드리 표.				
14	45.0	61.7	1	생건	_ 온확률(S)	*		= ()			
15	30.0	45.0	)	CHZ	동응답(U)		▲ 내용 표·	본( <u>R</u> )			
16	67.5	58.3	}		ミンロター	· [	레거시	대화 상자(L) 🕨	7 🏹	이제곱검정(	<u>C</u> )
17	32.5	46.7	'		s El Olano				0/1 0	항 <mark>(B)</mark>	
10	22.5	16 7			5 내립( <u>1</u> ) - 프립고				- 1998 년	(R)	
FILOIDI	비기(四) 변수 보기(	VA.		복합	갈 표 본( <u>L</u> )	*	***		· · · · · · · · · · · · · · · · · · ·	- 표본 K-S(1).	
GIUE	27(0)			間시험	s레이션					린 2-표보(2)	
대응 2-표	본(L)			품질	물 관리( <u>Q</u> )	•					
-					으 고 서 ^ ^			100		법 <u>K</u> -#폰(K)	
									EH	응 2-표본( <u>L</u> )	
									EH CH	응 K-표본( <u>S</u> )	)

분석(A) → 비모수검정(N) → 레거시 대화 상자(L) → 대응2-표본(L) 클릭

<ul> <li>✔ 사전점수</li> <li>✔ 사후점수</li> <li>✔ 사후점수</li> </ul>	검정 쌍(T): 대응(A) 변수1 변수2 1	★     정확(X)       ★     옵션(Q),       ★	통계량 ☑ 기술통계(D) ■ 사분위수(Q) 결측값 ◎ 검정별 결측값 제외(T) ◎ 목록별 결측값 제외(L)
<u>확인</u> 물여넣기(	김정 유형	도움말	계속 취소 도움말

- P2

사전점수와 사후점수를 "검정 쌍(T)"에 넣고 옵션 "기술통계" 클릭해도 되고 안해도 됩니다. 그리고 확인을 클릭하면 아웃풋이 나옵니다.

기술통계량

	Ν	평균	표준편차	최소값	최대값
사전점수	29	60.776	23.9987	10.0	95.0
사후점수	29	66.639	18.8777	40.0	95.0

# Wilcoxon 부호순위 검정

순위						
		Ν	평균순위	순위합		
사후점수 - 사전점수	음의 순위	6 ^a	17.83	107.00		
	양의 순위	22 ^b	13.59	299.00		
	등를	1 ^c				
	합계	29				

a. 사후점수 < 사전점수

b. 사후점수 > 사전점수

c. 사후점수 = 사전점수

검정 통계량^a



순위표를 보면, N칸의 음의 순위는 6개, 양의 순위는 22개, 동률은 1개로 표시되어 있 습니다. 이는 사전보다 사후에 점수가 떨어진 것은 음의 순위로, 사전보다 사후에 점수가 오른 것은 양의 순위로, 사전과 사후가 같으면 동률로 표시되는 것입니다.

양의 순위가 음의 순위보다 많은 것을 보더라도 사후점수가 높을 것이라는 것을 예측할 수 있습니다.

검정 통계량을 보시면,

근사 유의확률(양측)값이 .029로 유의수준 .05보다 작아 통계적으로 유의미하게 차이가 있는 것을 확인할 수 있었습니다. 3) Kruskal-Wallis H 검정



분석(A) → 비모수검정(N) → 레거시 대화 상자(L) → 독립K-표본(K) 클릭



검정변수에 종속변수를 넣고 집단변수에 "집단"을 넣고 "집단정의" 클릭한 후

최소값과 최대값으로 집단을 나누어 줍니다. 본 샘플은 집단이 4개라 1~4로 지정했습니 다.

Kruskal-Wallis의 H(K)을 체크하고 "확인"을 클릭하면 아웃풋이 나옵니다.

# Kruskal-Wallis 검정

순위					
	집단	Ν	평균순위		
점수	실험집단1	6	14.50		
	실험집단2	7	15.79		
	실험집단3	6	9.75		
	비교집단	4	5.00		
	합계	23			

검정 통계량^{a,b}

	점수	
카이제곱	7.945	
자유도	3	
근사 유의확률	.047	
a. Kruskal W 검정	allis	
b.집단변수:	집단	

평균순위가 높을수록 평균이 높습니다.

여기서는 실험집단2가 가장 평균순위가 높고 비교집단이 가장 평균순위가 낮습니다. 검정 통계량을 보시면,

근사 유의확률을 보면 됩니다. .048로 유의수준 .05보다 작아 통계적으로 유의미하게 차이가 있는 것을 확인할 수 있었습니다.

하지만 사후검증이 되지 않기 때문에 통계적으로 유의미한 차이가 있다는 것만 알 수 있습니다.

정규성과 표본수가 아닌 등분산성이 성립하지 않았을 때는 등분산을 가정하지 않는 사 후분석을 사용합니다. 이 방법은 분산분석(ANOVA)에서 자세히 다루었으니 참고하시면 됩니다.

## 지금까지 비모수통계 아웃풋을 보시면 아시겠지만, 비모수통계의 핵심은 연속형자료로 입 력된 데이터를 SPSS프로그램 자체에서 순위(서열척도)로 전환하여 분석한다는 것입니다.

### 예제) 비모수 통계분석

1) Mann-Whitney U 검정

실험집단(n=8)	비교집	단(n=8)		
평균순위 평균±표준편차 (순위합)	평균순위 (순위합)	평균±표준편차	Ζ	р
점수 10.94(87.50) 37.63±6.52	6.06(48.50)	30.63±3.16	-2.055	0.038

실험집단의 평균은 37.63점(평균순위는 10.94)로 비교집단 평균 30.63점(평균순위 6.06) 보다 높았으며, 이는 통계적으로 유의미한 차이가 있었다(p<.05).

2) Wilcoxon 부호-서열 검정

		Ν	평균순위	순위합		
사전-사후	Negative Ranks (음의 순위)	6 ^a	17.83	107.00	Z	р
	Positive Ranks (양의 순위)	22 ^b	13.59	229.00		
	Ties (동률)	$1^{c}$				
	Total (합계)	29				
a. 사후점수<사전점수						
b. 사후	·점수>사전점수				-2.187	0.029
c. 사후	점수=사전점수					

사전점수가 사후점수보다 큰 음의 순위는 6개였으며, 사후점수가 사전점수보다 큰 양의 순위는 22로 사전점수보다는 사후점수가 증가했으며, 이는 통계적으로 유의미한 차이가 있었다(p<.05).

### 3) Kruskal-Wallis H 검정

	집단	평균순위	평균±표준편차	$X^{2}$	р
점수	실험집단1	14.50	$76.67 \pm 6.86^{a}$		0.047
	실험집단2	15.79	78.43±7.11 ^ª	7.045	
	실험집단3	9.75	$70.83 \pm 9.58^{ab}$	7.940	0.047
	비교집단	5.00	$65.25 \pm 4.11^{b}$		

점수는 실험집단2가 평균 78.43점(평균순위 15.79)으로 가장 많았으며, 다음으로 실험집단 1 평균 76.67점(평균순위 14.50), 실험집단3 평균 70.83점(평균순위 9.75), 비교집단 평균 65.25점(평균순위 5.00)순이었으며, 이는 통계적으로 유의미한 차이가 있었다(p<.05). 사후 검증결과 실험집단1과 실험집단2가 비교집단보다 높은 것으로 나타났다.