

IBM SPSS Custom Tables
24

IBM

참고

이 정보와 이 정보가 지원하는 제품을 사용하기 전에, 반드시 113 페이지의 『주의사항』에 나오는 일반 정보를 읽으십시오.

제품 정보

이 개정판은 새 개정판에서 별도로 명시하지 않는 한, IBM SPSS Statistics의 버전 24, 릴리스 0. 수정사항 0 및 모든 후속 릴리스와 수정에 적용됩니다.

목차

제 1 장 테이블 작성기 인터페이스	1	중첩 테이블의 총계	39
테이블 작성기 인터페이스	1	레이어 변수 총계	40
테이블 작성	1	소개	41
테이블 작성	4	표시된 내용대로 소개 생성	41
누적변수	4	소개 범주 감추기	42
중첩 변수	4	레이어 변수 소개	43
레이어	5		
변수 이름 또는 레이블 보이기 및 감추기	6	제 5 장 범주형 변수의 계산된 범주	45
요약 통계	6	단순 계산된 범주	45
범주 및 총계	12	계산된 범주에서 범주 감추기	46
계산된 범주	15	계산된 범주의 소개 참조	47
공유 범주가 있는 변수에 대한 테이블 (Comperimeter 테이블)	16	계산된 범주를 사용하여 비전체적인 소개 표시	48
테이블 작성기 사용자 정의	16		
테이블 사용자 정의: 옵션 탭	17	제 6 장 공유 범주가 있는 변수 테이블	51
테이블 사용자 정의: 제목 탭	18	개수표	51
테이블 사용자 정의: 검정 통계량 탭	19	퍼센트 테이블	52
		총계 및 범주 제어	53
		공유 범주가 있는 테이블에 중첩	54
제 2 장 범주형 변수에 대한 단순 통계표	21	제 7 장 요약 통계	55
범주형 변수에 대한 단순 통계표	21	요약 통계 소스 변수	56
단일 범주형 변수	22	범주형 변수의 요약 통계 소스	56
퍼센트	22	척도변수의 요약 통계 소스	56
총계	23	누적변수	57
교차 분석표	23	범주형 변수의 사용자 정의 총계 요약 통계	58
교차 분석표의 퍼센트	24	범주 값 표시	59
표시 형식 제어	25		
주변 합계	25	제 8 장 척도변수 요약	61
범주 정렬 및 제어	26	척도변수 요약	61
		누적 척도변수	61
제 3 장 범주형 변수를 사용한 누적, 중첩 및 레이어 29		다중 요약 통계	62
범주형 변수 누적	29	개수, 유효수 및 결측	62
교차 분석표에서의 누적	29	변수마다 다른 요약	64
중첩 범주형 변수	30	범주별 그룹 요약	65
변수 레이블 출력없음	31	다중 그룹화 변수	65
중첩 교차 분석표	32	척도변수 내에 범주형 변수 중첩	66
레이어	34		
두 개의 누적 범주형 레이어 변수	35	제 9 장 신뢰구간	69
두 개의 중첩된 범주형 레이어 변수	35		
제 4 장 범주형 변수에 대한 총계 및 소개	37	제 10 장 검정 통계량	71
단일 변수의 단순 총계	37	검정 통계량	71
표시된 내용대로 총계 생성	38	독립성 검정(카이제곱)	71
총계의 출력 위치	38	독립성 검정에서 중첩과 누적이 효과	72
		열 평균 비교	74

열 평균 검정에서 중첩과 누적의 효과.	76	테이블에 결측값 포함	94
열 비율 비교.	77	제 13 장 표 형식 지정 및 사용자 정의	97
열 비율 검정에서 중첩과 누적의 효과.	81	표 형식 지정 및 사용자 정의	97
가중치와 다중 응답 세트에 대한 참고 사항.	83	요약 통계 표시 형식	97
제 11 장 다중 응답 세트	85	요약 통계에 대한 레이블 표시	99
개수, 반응, 퍼센트 및 총계	85	열 너비	100
다른 변수와 함께 다중 응답 세트 사용	88	빈 셀의 표시 값	100
통계량 소스 변수 및 사용 가능한 요약 통계	88	결측 통계의 표시 값	101
다중 범주 세트 및 중복 응답	89	제 14 장 표본 파일	103
다중 응답 세트에 대한 유의성 검증	90	주의사항	113
다중 응답 세트에 대한 독립성 검정	90	상표	115
다중 응답 세트에 대한 열 평균 비교.	91	색인	117
제 12 장 결측값	93		
결측값이 없는 테이블	93		

제 1 장 테이블 작성기 인터페이스

테이블 작성기 인터페이스

테이블 사용자 정의는 단순 끌어서 놓기 테이블 작성기 인터페이스를 통해 변수와 옵션을 선택하기만 해도 테이블을 미리 볼 수 있습니다. 테이블 사용자 정의는 또한 창의 크기 변경이라든지 분할창에서 영역 크기 변경 기능 등 기존의 대화 상자에서 볼 수 없는 유연성을 제공합니다.

테이블 작성

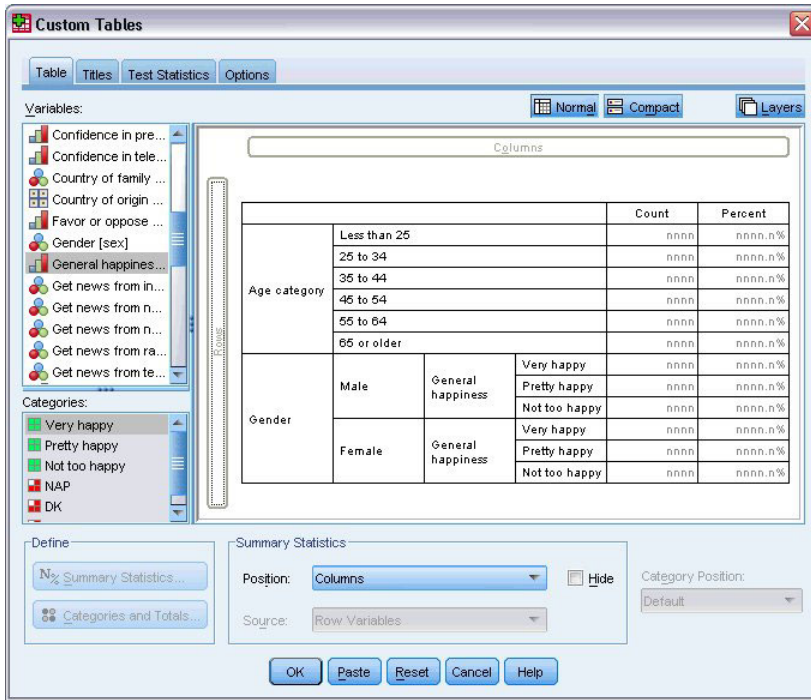


그림 1. 테이블 사용자 정의 대화 상자, 테이블 탭

테이블 작성기의 테이블 탭에서 테이블에 표시할 변수와 요약 측도를 선택합니다.

변수 목록. 데이터 파일의 변수는 분할창의 왼쪽 상단 영역에 표시됩니다. 테이블 사용자 정의에서는 변수에 대한 두 측정 수준을 구분하여 해당 측정 수준에 따라 서로 다르게 처리합니다.

범주형. 고유한 값/범주 수가 제한된 데이터(예: 성별 또는 종교)입니다. 문자열(문자숫자형) 변수 또는 숫자 코드를 사용하여 범주를 나타내는 숫자변수(예: 0 = 남성, 1 = 여성)가 범주형 변수에 해당될 수 있습니다. 질적 데이터라고도 합니다. 범주형 변수는 명목이거나 순서일 수 있습니다.

- **명목(Nominal)**. 변수의 값이 고유한 순위가 없는 범주를 나타내는 경우 해당 변수는 명목으로 취급될 수 있습니다. 예를 들어, 직원이 근무하는 회사의 부서가 있습니다. 종교, 우편번호 또는 종교 단체 등이 명목 변수에 해당합니다.
- **순서(Ordinal)**. 변수의 값이 고유한 순위가 있는 범주를 나타내는 경우 해당 변수는 순서로 취급될 수 있습니다. 예를 들어, 매우 불만족에서 매우 만족에 이르는 서비스 만족도 수준이 있습니다. 순서변수의 예로는 만족도나 신뢰도를 나타내는 태도 스코어 및 선호도 등급 스코어가 있습니다.

척도. 구간 또는 비율 척도로 측정된 데이터에서 데이터 값은 값의 순서와 차이를 모두 나타냅니다. 예를 들어, \$72,195의 급여는 \$52,398보다 높은 순위이고 두 값의 차이는 \$19,797입니다. 이러한 데이터는 양적, 연속형 데이터라고도 합니다.

범주형 변수는 테이블의 범주(행, 열, 레이어)를 정의하고, 기본값 요약 통계는 개수(각 범주 내의 케이스 수)입니다. 예를 들어, 범주형 성별 변수의 기본값 테이블에는 단순히 남자 수와 여자 수가 표시됩니다.

척도변수는 일반적으로 범주형 변수의 범주 내에 요약되고 기본값 요약 통계는 평균입니다. 예를 들어, 성별 범주 내의 수입 기본값 테이블에는 남자의 평균 수입과 여자의 평균 수입이 표시됩니다.

범주형 변수를 사용하여 그룹을 정의하지 않고 척도변수를 직접 요약할 수도 있습니다. 그렇게 하면 다중 척도 변수의 요약 누적에 유용합니다. 자세한 정보는 4 페이지의 『누적변수』 주제를 참조하십시오.

다중 응답 세트

테이블 사용자 정의에서는 **다중 응답 세트**라는 특수 "변수"도 지원합니다. 다중 응답 세트는 실제로 일반적인 의미의 "변수"는 아닙니다. 다중 응답 세트는 데이터 편집기에서 볼 수 없으며 다른 프로시저에서 이 변수를 인식할 수도 없습니다. 다중 응답 세트는 반응자가 두 개 이상의 대담을 선택할 수 있는 질문에 대해 다중 변수를 사용하여 반응을 기록합니다. 다중 응답 세트는 범주형 변수처럼 처리되며 범주형 변수로 수행할 수 있는 대부분의 작업은 다중 응답 세트로도 수행할 수 있습니다. 자세한 정보는 85 페이지의 제 11 장 『다중 응답 세트』 주제를 참조하십시오.

변수 목록의 각 변수 옆에 있는 아이콘은 변수 유형을 식별합니다.

표 1. 측정 수준 아이콘














	숫자	문자열	날짜	시간
척도(연속형)		해당 없음		
순서				
명목				

표2. 다중 응답 세트 아이콘

다중 응답 세트 유형	아이콘
다중 응답 세트, 다중 범주	
다중 응답 세트, 다중 이분형	

변수 목록에서 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주형 또는 척도를 선택하여 테이블 작성기에서 변수의 측정 수준을 변경할 수 있습니다. 데이터 편집기의 변수 보기에서 변수의 측정 수준을 변경할 수 있습니다. 명목 또는 순서로 정의된 변수는 테이블 사용자 정의에서 범주형으로 취급됩니다.

범주. 변수 목록에서 범주형 변수를 선택하면 해당 변수에 대해 정의된 범주가 범주 목록에 표시됩니다. 이러한 범주는 테이블에서 변수를 사용할 때 캔버스 분할창에도 표시됩니다. 변수에 정의된 범주가 없는 경우 변수 목록과 캔버스 분할창에 플레이스홀더 범주 1과 범주 2를 표시합니다.

테이블 작성기에 표시되는 정의된 범주는 서로 다른 데이터 값에 할당된 설명 레이블(예: 숫자값 0과 1은 값 레이블에 남자 및 여자로 표시)인 값 레이블을 기준으로 합니다. 값 레이블은 데이터 편집기의 변수 보기에서 정의하거나 데이터 편집기 창의 데이터 메뉴에서 변수 특성 정의를 사용하여 정의할 수 있습니다.

캔버스 분할창. 변수를 캔버스 분할창의 행과 열로 끌어다 놓고 테이블을 작성합니다. 캔버스 분할창에는 작성할 테이블의 미리보기가 표시됩니다. 캔버스 분할창에는 셀에 있는 실제 데이터 값은 표시되지 않지만 최종 테이블의 레이아웃에 대한 매우 정확한 보기가 표시됩니다. 범주형 변수의 경우 데이터 파일에 값 레이블이 정의되지 않은 고유한 값이 포함되어 있다면 실제 테이블에는 미리보기보다 더 많은 범주가 포함될 수 있습니다.

- 보통 보기에는 범주형 변수의 범주 및 요약 통계에 대한 행 또는 열을 포함하여 테이블에 포함되는 모든 행과 열이 표시됩니다.
- 최소 보기에는 테이블에 포함되는 행과 열에 대한 미리보기 없이 테이블에 있는 변수만 표시됩니다.

테이블 작성을 위한 기본 규칙과 제한

- 범주형 변수의 경우 요약 통계는 통계 소스 차원에서 가장 안쪽에 있는 변수를 기준으로 합니다.
- 범주형 변수에 대한 기본값 통계 소스 차원(행 또는 열)은 캔버스 분할창으로 변수를 끌어다 놓는 순서를 기준으로 합니다. 예를 들어, 변수를 행 트레이로 먼저 끌어다 놓으면 행 차원이 기본값 통계 소스 차원이 됩니다.
- 척도변수는 행 차원이나 열 차원에서 가장 안쪽에 있는 변수의 범주 내에만 요약될 수 있습니다. 척도변수는 테이블의 모든 수준에 배치될 수 있지만 가장 안쪽에 있는 수준에서 요약됩니다.
- 척도변수는 다른 척도변수 내에 요약될 수 없습니다. 다중 척도변수에 대한 요약을 누적하거나 범주형 변수의 범주 내에 척도변수를 요약할 수 있습니다. 척도변수를 다른 척도변수 내에 중첩할 수 없습니다. 또한 한 척도변수는 행 차원에 넣고 다른 척도변수는 열 차원에 넣을 수 없습니다.
- 활성 데이터 세트의 어느 변수가 정의된 값 레이블을 12000개 이상 포함할 경우 테이블 작성기를 사용하여 테이블을 만들 수 없습니다. 표에서 이 한계를 초과하는 변수를 포함시킬 필요가 없는 경우 그러한 변수를

제외한 변수 세트를 정의하고 적용할 수 있습니다. 정의된 값 레이블을 12000개 이상 갖고 있는 어느 변수를 포함시켜야 할 경우 CTABLES 명령 구문을 사용하여 테이블을 생성할 수 있습니다.

테이블 작성

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 하나 이상의 변수를 캔버스 분할창의 행 또는 열 영역으로 끌어다 놓습니다.
3. 확인을 클릭하여 테이블을 작성합니다.
4. 캔버스 분할창에서 변수를 선택합니다.
5. 변수를 캔버스 분할창 바깥쪽으로 끌거나 Delete 키를 누릅니다.

변수의 측정 수준을 변경하려면 다음과 같이 하십시오.

6. 변수 목록에서 변수를 마우스 오른쪽 단추로 클릭합니다. 이 작업은 변수 목록에서만 가능하고 캔버스에서는 가능하지 않습니다.
7. 팝업 메뉴에서 범주형 또는 척도를 선택합니다.

누적변수

누적은 개별 테이블을 취하여 동일한 출력 화면에 붙여넣는 것입니다. 예를 들어, 같은 테이블에서 성별과 연령범주에 대한 정보를 별도의 섹션으로 표시할 수 있습니다.

변수 누적

1. 변수 목록에서 누적할 모든 변수를 선택한 다음 캔버스 분할창의 행 또는 열로 끌어다 놓습니다.

또는

2. 개별 변수를 행에서 기존 변수 위나 아래 또는 열에서 기존 변수 오른쪽이나 왼쪽으로 끌어다 놓습니다.

표 3. 누적 범주형 변수.

변수	범주	요약 통계
변수 1	범주 1	123
	범주 2	456
변수 2	범주 1	123
	범주 2	456
	범주 3	789

자세한 정보는 29 페이지의 『범주형 변수 누적』 주제를 참조하십시오.

중첩 변수

두 범주형 변수 간의 관계를 표시할 수 있다는 점에서는 중첩은 교차 분석표와 비슷하지만, 교차 분석표와 달리 중첩에서는 한 변수가 같은 차원의 다른 변수 내에 중첩됩니다. 예를 들어, 행 차원의 연령범주 내에 성별을 중첩하여 각 연령 범주에 대한 여자 및 남자의 수를 표시할 수 있습니다.

범주형 변수 내에 척도변수를 중첩할 수도 있습니다. 예를 들어, 성별 내에 수입을 중첩하여 남자와 여자의 개별 평균(또는 중위수 또는 기타 요약 척도) 수입 값을 표시할 수 있습니다.

중첩 변수

1. 범주형 변수를 캔버스 분할창의 행 또는 열 영역으로 끌어다 놓습니다.
2. 범주형 변수 또는 척도변수를 범주형 행 변수의 왼쪽이나 오른쪽 또는 범주형 열 변수의 위나 아래로 끌어다 놓습니다.

표 4. 중첩 범주형 변수.

변수 1	변수 2	요약 통계
범주 1	범주 1	12
	범주 2	34
	범주 3	56
범주 2	범주 1	12
	범주 2	34
	범주 3	56

자세한 정보는 30 페이지의 『중첩 범주형 변수』 주제를 참조하십시오.

참고: 테이블 사용자 정의는 레이어 분할 파일 처리를 사용하지 않습니다. 레이어 분할 파일과 같은 결과를 얻으려면 분할 파일 변수를 테이블의 가장 바깥쪽에 있는 중첩 레이어에 두십시오.

레이어

레이어를 사용하면 테이블에 차원 깊이를 추가하여 3차원 "큐브"를 만들 수 있습니다. 레이어는 중첩 또는 누적과 유사하지만 레이어 범주는 한 번에 하나만 표시할 수 있다는 점에서 다릅니다. 예를 들어, 행 변수로 연령범주를 사용하고 레이어 변수로 성별을 사용하면 여자와 남자에 대한 정보가 다른 레이어에 표시되는 테이블을 작성할 수 있습니다.

레이어 작성

1. 테이블 작성기의 테이블 탭에서 레이어를 클릭하여 레이어 목록을 표시합니다.
2. 레이어를 정의하는 척도변수 또는 범주형 변수를 레이어 목록으로 끌어다 놓습니다.

레이어 목록에 척도변수와 범주형 변수를 혼합할 수 없습니다. 모든 변수는 같은 유형이어야 합니다. 다중 응답 세트는 레이어 목록에서 범주형 변수로 취급됩니다. 레이어의 척도변수는 항상 누적됩니다.

다중 범주형 레이어 변수가 있는 경우 레이어를 누적하거나 중첩할 수 있습니다.

- 각 범주를 레이어로 표시는 누적과 동일합니다. 각 레이어 변수의 각 범주에 대해 별도의 레이어가 표시됩니다. 총 레이어 수는 각 레이어 변수의 범주 수에 대한 단순 총계입니다. 예를 들어, 세 개의 레이어 변수가 있고 각 변수에 세 개씩의 범주가 있는 경우 테이블에는 아홉 개의 레이어가 포함됩니다.
- 각 범주 조합을 레이어로 표시는 레이어 중첩 또는 교차 분석표와 동일합니다. 총 레이어 수는 각 레이어 변수에 대한 범주 수의 곱입니다. 예를 들어, 세 개의 레이어 변수가 있고 각 변수에 범주가 세 개씩 있는 경우 테이블에는 27개의 레이어가 포함됩니다.

변수 이름 또는 레이블 보이기 및 감추기

다음 옵션을 변수 이름 및 레이블 표시에 사용할 수 있습니다.

- 변수 레이블만 표시. 정의된 변수 레이블이 없는 변수인 경우 변수 이름이 표시됩니다. 기본값 설정입니다.
- 변수 이름만 표시
- 변수 레이블과 변수 이름 모두 표시
- 변수 이름 또는 변수 레이블 표시 안 함. 변수 레이블이나 이름이 포함된 열/행이 캔버스 분할창의 테이블 미리보기에 표시되더라도 이 열/행은 실제 테이블에는 표시되지 않습니다.

변수 레이블 또는 변수 이름을 보이거나 감추려면 다음과 같이 하십시오.

1. 캔버스 분할창의 테이블 미리보기에서 변수를 마우스 오른쪽 단추로 클릭합니다.
2. 팝업 메뉴에서 변수 레이블 보이기 또는 변수 이름 보이기를 선택하여 레이블 또는 이름 표시를 설정하거나 해제합니다. 선택 옆의 체크 표시는 해당 항목이 표시됨을 나타냅니다.

요약 통계

요약 통계 대화 상자에서 수행할 수 있는 작업은 다음과 같습니다.

- 테이블에서 요약 통계 추가 및 제거
- 통계에 대한 레이블 바꾸기
- 통계의 순서 바꾸기
- 소수점 이하 자릿수를 포함하여 통계의 형식 바꾸기

여기서 사용할 수 있는 요약 통계 및 기타 옵션은 대화 상자의 상단에 표시된 대로 요약 통계 소스 변수의 측정 수준에 따라 다릅니다. 요약 통계의 소스(요약 통계 기준 변수)는 다음에 의해 결정됩니다.

- 측정 수준. 테이블 또는 누적표의 테이블 섹션에 척도변수가 있는 경우 요약 통계는 척도변수를 기준으로 합니다.
- 변수 선택 순서. 범주형 변수에 대한 기본값 통계 소스 차원(행 또는 열)은 캔버스 분할창으로 변수를 끌어서 놓는 순서를 기준으로 합니다. 예를 들어, 변수를 행 영역으로 먼저 끌면 행 차원이 기본값 통계 소스 차원이 됩니다.
- 중첩. 범주형 변수의 경우 요약 통계는 통계 소스 차원에서 가장 안쪽에 있는 변수를 기준으로 합니다.

누적표에는 다중 요약 통계 소스 변수(척도변수와 범주형 변수 모두)가 있을 수 있지만 각 테이블 섹션에는 요약 통계 소스가 하나만 있습니다.

요약 통계 소스 차원 바꾸기

1. 테이블 탭의 요약 통계에 있는 소스 드롭다운 목록에서 차원(행, 열 또는 레이어)을 선택합니다.

테이블에 표시된 요약 통계 제어

1. 테이블 탭의 캔버스 분할창에서 요약 통계 소스 변수를 선택합니다.
2. 테이블 탭의 정의에서 요약 통계를 클릭합니다.

또는

3. 캔버스 분할창에서 요약 통계 소스 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
4. 테이블에 포함할 요약 통계를 선택합니다. 화살표를 사용하여 선택한 통계를 통계 목록에서 출력 목록으로 이동할 수도 있고 선택한 통계를 통계 목록에서 출력 목록으로 끌어다 놓을 수도 있습니다.
5. 위쪽 또는 아래쪽 화살표를 클릭하여 현재 선택한 요약 통계의 출력 위치를 변경합니다.
6. 형식 드롭다운 목록에서 선택한 요약 통계에 대한 출력 형식을 선택합니다.
7. 선택한 요약 통계의 소수점이하자리 셀에 표시할 소수점 이하 자릿수를 입력합니다.
8. 선택한 항목에 적용을 클릭하여 캔버스 분할창에서 현재 선택한 변수에 대해 선택한 요약 통계를 포함합니다.
9. 모든 항목에 적용을 클릭하면 캔버스 분할창에 있는 동일한 유형의 모든 변수 누적에 대해 선택한 요약 통계가 포함됩니다.

참고: 모든 항목에 적용은 캔버스 분할창에 이미 있는 동일한 유형의 변수 누적에 대해서만 선택한 항목에 적용과 다르게 적용됩니다. 두 경우 모두 선택한 요약 통계가 테이블에 추가된 동일한 유형의 추가 변수 누적에 자동으로 포함됩니다.

범주형 변수의 요약 통계

범주형 변수에 사용 가능한 기초적 통계는 개수와 퍼센트입니다. 총계 및 소계에 대한 사용자 정의 요약 통계를 지정할 수도 있습니다. 이러한 사용자 정의 요약 통계에는 일부 순서 범주형 변수에 적합한 중심 경향 측도(예: 평균 및 중위수)와 산포(예: 표준 편차)가 포함됩니다. 자세한 정보는 10 페이지의 『범주형 변수의 사용자 정의 총계 요약 통계』 주제를 참조하십시오.

개수. 테이블의 각 셀에 있는 케이스 수 또는 다중 응답 세트의 반응 수입니다. 가중치가 적용되면 이 값은 가중치를 부여한 개수입니다.

- 가중치가 적용되면 이 값은 가중치를 부여한 개수입니다.
- 가중치를 부여한 개수는 글로벌 데이터세트 가중치(데이터 > 케이스 가중치)와 사용자 정의 테이블 대화 상자의 옵션 탭에 지정된 유효 기준 가중치 변수를 사용한 가중치에서 모두 동일합니다.

가중치를 부여하지 않은 개수. 테이블의 각 셀에 있는 가중치를 부여하지 않은 케이스 수입니다. 가중치를 부여하지 않은 개수는 가중치가 적용되는 경우에만 개수와 다릅니다.

조정된 개수. 조정된 계수는 유효 기준 가중치 계산에 사용됩니다. 유효 기준 가중치 변수(옵션 탭)를 사용하지 않는 경우 조정된 개수는 개수와 동일합니다.

열 퍼센트. 각 열 내의 퍼센트입니다. 하위 테이블에서 각 열의 퍼센트(단순 퍼센트의 경우) 합계는 100%입니다. 열 퍼센트는 일반적으로 범주형 행 변수가 있는 경우에만 유용합니다.

행 퍼센트. 각 행 내의 퍼센트입니다. 하위 테이블에서 각 행의 퍼센트(단순 퍼센트의 경우) 합계는 100%입니다. 행 퍼센트는 일반적으로 범주형 열 변수가 있는 경우에만 유용합니다.

레이어 행 및 레이어 열 퍼센트. 중첩된 테이블에서 모든 하위 테이블의 행 또는 열 퍼센트(단순 퍼센트의 경우) 합계는 100%입니다. 테이블에 레이어가 포함되어 있는 경우 각 레이어에서 중첩된 모든 하위 테이블의 행 또는 열 퍼센트 총계는 100%입니다.

레이어 퍼센트. 각 레이어 내의 퍼센트입니다. 단순 퍼센트의 경우 현재 표시된 레이어 내의 셀 퍼센트 총계가 100%입니다. 레이어 변수가 없는 경우 레이어 퍼센트는 표 퍼센트와 동일합니다.

테이블 퍼센트. 각 셀의 퍼센트는 전체 테이블을 기준으로 합니다. 모든 셀 퍼센트는 동일한 총 케이스 수를 기준으로 하며 전체 테이블에서 총계가 100%(단순 퍼센트의 경우)입니다.

하위 테이블 퍼센트. 각 셀의 퍼센트는 하위 테이블을 기준으로 합니다. 하위 테이블의 모든 셀 퍼센트는 동일한 총 케이스 수를 기준으로 하며 하위 테이블 내에서 총계가 100%입니다. 중첩된 테이블의 경우 가장 안쪽에 있는 중첩 수준의 앞에 오는 변수가 하위 테이블을 정의합니다. 예를 들어, 연령범주 내에 성별, 성별 범주 내에 결혼상태 범주가 중첩된 통계표에서는 성별이 하위 테이블을 정의합니다.

신뢰구간

- 하위 및 상위 신뢰 한계는 개수, 백분율, 평균, 중앙값, 백분위수 및 합계에 사용할 수 있습니다.
- 레이블의 텍스트 문자열 "&[신뢰수준]"은 테이블의 열 레이블에 신뢰수준을 포함합니다.
- 표준 오류는 개수, 백분율, 평균 및 합계에 사용할 수 있습니다.
- 신뢰구간 및 표준 오류는 다중 응답 세트에 사용할 수 없습니다.

수준 백분율로 표현되는 신뢰구간의 신뢰수준입니다. 이 값은 0보다 크고 100보다 작아야 합니다.

다중 응답 세트

다중 응답 세트에는 케이스, 반응, 개수 등을 기준으로 하는 퍼센트가 포함될 수 있습니다. 자세한 정보는 9 페이지의 『다중 응답 세트에 대한 요약 통계』 주제를 참조하십시오.

누적표

퍼센트 계산에서는 누적변수로 정의된 테이블의 각 섹션이 별도의 테이블로 취급됩니다. 누적표의 각 섹션에서 레이어 행, 레이어 열 및 표 퍼센트의 합계는 100%(단순 퍼센트의 경우)입니다. 다른 퍼센트 계산의 퍼센트 기준은 각 누적표 섹션의 케이스를 기준으로 합니다.

퍼센트 기준

퍼센트는 계산 기준의 결측값 처리에 따라 세 가지 방법으로 계산할 수 있습니다.

단순 퍼센트. 퍼센트는 테이블에 사용된 케이스 수를 기준으로 하며 합계는 항상 100%입니다. 범주를 테이블에서 제외하면 해당 범주의 케이스가 기준에서 제외됩니다. 시스템-결측값이 있는 케이스는 기준에서 항상 제외됩니다. 사용자 결측값을 가지는 케이스는 사용자 결측 범주가 테이블에서 제외되는 경우(기본값) 제외되고 사용자 결측 범주가 테이블에 포함되는 경우 포함됩니다. 이름에 유효수 또는 합계 N 이 없는 퍼센트는 단순 퍼센트입니다.

합계 N 퍼센트. 시스템-결측값 및 사용자 결측값을 가지는 케이스는 단순 퍼센트 기준에 추가됩니다. 퍼센트 합계가 100%보다 작을 수 있습니다.

유효수 퍼센트. 사용자 결측값을 가지는 케이스는 사용자 결측 범주가 테이블에 포함되더라도 단순 퍼센트 기준에서 제거됩니다.

참고: 사용자 결측 범주 외에 수동으로 제외된 범주에 있는 케이스는 기준에서 항상 제외됩니다.

다중 응답 세트에 대한 요약 통계

다음 추가 요약 통계를 다중 응답 세트에 사용할 수 있습니다.

열/행/레이어-반응 %. 반응을 기준으로 하는 퍼센트입니다.

열/행/레이어 반응 % (기준: 개수). 반응은 분자이고 총계는 분모입니다.

열/행/레이어 개수 % (기준: 반응). 개수는 분자이고 총 반응 수는 분모입니다.

레이어 열/행 반응 %. 하위 테이블 간 퍼센트입니다. 반응을 기준으로 하는 퍼센트입니다.

레이어 열/행 반응 % (기준: 개수). 하위 테이블 간 퍼센트입니다. 반응은 분자이고 총계는 분모입니다.

레이어 열/행 반응 % (기준: 반응). 하위 테이블 간 퍼센트입니다. 개수는 분자이고 총 반응 수는 분모입니다.

반응. 반응 개수입니다.

하위 표/표 반응 %. 반응을 기준으로 하는 퍼센트입니다.

하위 표/표 반응 % (기준: 개수). 반응은 분자이고 총계는 분모입니다.

하위 표/표 개수 % (기준: 반응). 개수는 분자이고 총 반응 수는 분모입니다.

척도변수 및 범주형 사용자 정의 총계의 요약 통계

범주형 변수에 사용할 수 있는 개수 및 퍼센트 뿐만 아니라 다음 요약 통계는 척도변수에도 사용할 수 있으며 범주형 변수의 사용자 정의 총계 및 소계 요약으로도 사용할 수 있습니다. 이러한 요약 통계는 다중 응답 세트나 문자열(문자숫자형) 변수에는 사용할 수 없습니다.

평균. 산술 평균으로 합계를 케이스 수로 나눈 값입니다.

중위수. 전체 케이스의 절반이 위 아래에 해당되는 값으로 제50 백분위수입니다.

모드. 가장 자주 나오는 값입니다. 동률이 있는 경우 가장 작은 값이 표시됩니다.

최소값. 가장 작은(가장 낮은) 값입니다.

최대값. 가장 큰(가장 높은) 값입니다.

결측. 결측값의 개수(사용자 및 시스템-결측값)입니다.

백분위수. 제5, 25, 75, 95 또는 99 백분위수를 포함할 수 있습니다.

범위. 최대값과 최소값의 차이입니다.

표준 편차. 평균에 대한 산포 측도입니다. 정규 분포에서 케이스의 68%는 평균의 표준 편차 내에 있으며 케이스의 95%는 2배 표준 편차 내에 있습니다. 예를 들어, 평균 연령이 45세이고 표준 편차가 10인 경우 정규 분포(분산의 제곱근) 내에서 케이스의 95%는 25세와 65세 사이에 있습니다.

합계. 값의 합계입니다.

합계 퍼센트. 합계를 기준으로 하는 퍼센트입니다. 행과 열(하위 테이블 내), 전체 행과 열(하위 테이블 간), 레이어, 하위 표, 전체 테이블 등에 사용할 수 있습니다.

합계 N. 비결측값, 사용자 결측값 및 시스템-결측값의 개수입니다. 사용자 결측 범주가 아닌 수동으로 제외된 범주에는 케이스가 포함되지 않습니다.

조정된 합계 N. 조정된 합계 N은 유효 기준 가중치 계산에 사용됩니다. 유효 기준 가중치 변수(옵션 탭)를 사용하지 않는 경우 조정된 합계 N은 합계 N과 동일합니다. 이 통계는 다중 응답 세트에 사용할 수 없습니다.

유효수. 비결측 값의 개수. 사용자 결측 범주가 아닌 수동으로 제외된 범주에는 케이스가 포함되지 않습니다.

조정된 유효수. 조정된 유효수는 유효 기준 가중치 계산에 사용됩니다. 유효 기준 가중치 변수(옵션 탭)를 사용하지 않는 경우 조정된 유효수는 유효수와 동일합니다. 이 통계는 다중 응답 세트에 사용할 수 없습니다.

분산. 평균에 대한 산포 측도값으로, 평균으로부터의 제곱 편차 합계를 케이스 수에서 1을 뺀 값으로 나눈 값과 같습니다. 분산은 변수 자체의 제곱(표준 편차의 제곱) 단위로 측도됩니다.

신뢰구간

- 하위 및 상위 신뢰 한계는 개수, 백분율, 평균, 중앙값, 백분위수 및 합계에 사용할 수 있습니다.
- 레이블의 텍스트 문자열 "&[신뢰수준]"은 테이블의 열 레이블에 신뢰수준을 포함합니다.
- 표준 오류는 개수, 백분율, 평균 및 합계에 사용할 수 있습니다.
- 신뢰구간 및 표준 오류는 다중 응답 세트에 사용할 수 없습니다.

수준 백분율로 표현되는 신뢰구간의 신뢰수준입니다. 이 값은 0보다 크고 100보다 작아야 합니다.

누적표

누적변수로 정의한 각 테이블 섹션이 별도의 테이블로 취급되고 이에 따라 요약 통계가 계산됩니다.

범주형 변수의 사용자 정의 총계 요약 통계

총계나 소계가 포함된 범주형 변수 테이블의 경우 각 범주에 대해 표시된 요약과 다르게 요약 통계를 표시할 수 있습니다. 예를 들어, 순서 범주형 행 변수에 대해서는 개수 및 열 퍼센트를 표시하고 "총계" 통계에 대해서는 중위수를 표시할 수 있습니다.

사용자 정의 총계 요약 통계를 사용하여 범주형 변수에 대한 테이블을 작성하려면 다음과 같이 하십시오.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

테이블 작성기가 열립니다.

2. 범주형 변수를 캔버스의 행 또는 열 영역으로 끌어다 놓습니다.
3. 캔버스에서 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
4. 총계 확인 상자를 선택한 다음 적용을 클릭합니다.
5. 캔버스에서 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
6. 총계 및 소계에 대한 사용자 정의 요약 통계 확인란을 선택한 다음 원하는 사용자 정의 요약 통계를 선택합니다.

기본적으로 사용자 정의 요약을 포함한 모든 요약 통계는 범주형 변수를 포함하는 차원과 반대 차원에 표시됩니다. 예를 들어, 범주형 행 변수가 있을 경우 요약 통계는 다음과 같이 테이블의 열을 정의합니다.

표 5. 행별 순서 변수 범주, 열별 요약 통계 개수 및 평균.

변수	범주	개수	평균
변수 1	찬성 1	196	2.29
	중립 2	936	
	반대 3	744	
	총계	1876	

범주형 변수와 동일한 차원에 요약 통계를 표시하려면 다음과 같이 하십시오.

7. 테이블 작성기의 테이블 탭에 있는 요약 통계의 위치 드롭다운 목록에서 차원을 선택합니다.

예를 들어, 범주형 변수가 행에 표시되는 경우 드롭다운 목록에서 행을 선택합니다.

요약 통계 표시 형식

다음 표시 형식 옵션을 사용할 수 있습니다.

nnnn. 단순 숫자입니다.

nnnn%. 퍼센트 기호를 값의 끝에 추가합니다.

자동. 소수점 이하 자릿수를 포함하여 정의된 변수 표시 형식입니다.

N=nnnn. 값 앞에 N=을 표시합니다. 요약 통계 레이블이 표시되지 않는 테이블의 개수, 유효수 N, 총계 N 등에 유용합니다.

(nnnn). 모든 값을 괄호로 묶습니다.

(nnnn)(음수 값). 음수 값만 괄호로 묶습니다.

(nnnn%). 모든 값을 괄호로 묶고 값의 끝에 퍼센트 기호를 붙입니다.

n,nnn.n. 쉼표 형식. 로케일 설정에 관계 없이 콤마가 그룹 구분 문자로 사용되고 마침표가 소수점 구분자로 사용됩니다.

n.nnn,n. 점 형식. 로케일 설정에 관계 없이 마침표가 그룹 구분 문자로 사용되고 콤마가 소수점 구분자로 사용됩니다.

\$n,nnn.n. 달러 형식. 달러 기호가 값 앞에 표시되며 로케일 설정에 관계 없이 콤마가 그룹 구분 문자로 사용되고 마침표가 소수점 구분자로 사용됩니다.

CCA, CCB, CCC, CCD, CCE. 사용자 정의 통화 형식입니다. 각 사용자 정의 통화에 대해 현재 정의된 형식이 목록에 표시됩니다. 이러한 형식은 옵션 대화 상자(편집 메뉴, 옵션)의 통화 탭에 정의됩니다.

일반 규칙 및 제한

- 자동을 제외하고 소수점 이하 자릿수는 소수점이하자리 열 설정에 따라 결정됩니다.
- 콤마, 달러 및 점 형식을 제외하고 Windows 국가별 옵션 제어판의 현재 로케일에 정의된 소수점 구분자가 사용됩니다.
- 콤마/달러 및 점 형식은 그룹 구분 문자로 각각 콤마 또는 마침표를 표시하지만 작성할 때 현재 로케일 설정(Windows 국가별 옵션 제어판)을 기준으로 그룹 구분 문자를 표시하는 데 사용할 수 있는 표시 형식이 없습니다.

범주 및 총계

범주 및 총계 대화 상자에서는 다음 작업을 수행할 수 있습니다.

- 범주 순서 바꾸기 및 제외
- 소계 및 총계 삽입
- 계산된 범주 삽입
- 빈 범주 포함 또는 제외
- 결측값을 포함하도록 정의된 범주 포함 또는 제외
- 값 레이블이 정의되지 않은 범주 포함 또는 제외
- 이 대화 상자는 범주형 변수와 다중 응답 세트에만 사용할 수 있으며 척도변수에는 사용할 수 없습니다.
- 범주가 서로 다른 다중 변수를 선택한 경우 소계를 삽입하거나, 계산된 범주를 삽입하거나, 범주를 제외하거나, 범주 순서를 수동으로 바꿀 수 없습니다. 이러한 작업은 캔버스 미리보기에서 다중 변수를 선택하고 선택한 모든 변수에 대해 이 대화 상자를 동시에 액세스하는 경우에만 가능합니다. 이러한 작업을 각 변수에 대해 개별적으로 수행하는 것은 가능합니다.
- 값 레이블을 정의하지 않은 변수의 경우 범주 정렬 및 총계 삽입만 수행할 수 있습니다.

범주 및 총계 대화 상자 액세스

1. 범주형 변수 또는 다중 응답 세트를 캔버스 분할창으로 끌어다 놓습니다.
2. 캔버스 분할창에서 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.

또는

3. 캔버스 분할창에서 변수를 선택하고 테이블 탭의 정의에서 범주 및 총계를 클릭합니다.

캔버스 분할창의 동일한 차원에서 다중 범주형 변수를 선택할 수도 있습니다.

4. 캔버스 분할창에서 각 변수를 Ctrl 키와 함께 클릭합니다.

또는

5. 캔버스 분할창에서 테이블 미리보기의 바깥쪽을 클릭하고 선택할 변수가 포함된 영역을 클릭한 다음 끌어서 선택합니다.

또는

6. 차원에서 임의의 변수를 마우스 오른쪽 단추로 클릭하고 모든 [차원] 변수 선택을 선택하여 해당 차원의 모든 변수를 선택합니다.

범주 순서 바꾸기

범주의 순서를 수동으로 바꾸려면 다음과 같이 하십시오.

1. 목록에서 범주를 선택합니다.
2. 위쪽 또는 아래쪽 화살표를 클릭하여 목록에서 범주를 위쪽 또는 아래쪽으로 이동합니다.

또는

3. 범주에 대한 값 열을 클릭하고 다른 차원으로 끌어다 놓습니다.

범주 제외

1. 목록에서 범주를 선택합니다.
2. 제외 목록 옆의 화살표를 클릭합니다.

또는

3. 범주에 대한 값 열을 클릭하고 목록 바깥으로 끌어다 놓습니다.

범주를 제외하면 정의된 값 레이블이 없는 범주도 함께 제외됩니다.

범주 정렬

데이터 값, 값 레이블, 셀 개수 또는 요약 통계를 기준으로 범주를 오름차순 또는 내림차순으로 정렬할 수 있습니다.

1. 범주 정렬에서 기준 드롭다운 목록을 클릭하고 사용할 정렬 기준인 값, 레이블, 개수 또는 요약 통계(예: 평균, 중위수 또는 최빈값)를 선택합니다. 정렬에 사용할 수 있는 요약 통계는 테이블에 표시할 요약 통계에 따라 다릅니다.
2. 순서 드롭다운 목록을 클릭하여 정렬 순서(오름차순 또는 내림차순)를 선택합니다.

범주를 제외한 경우에는 범주 정렬을 사용할 수 없습니다.

소계

1. 목록에서 소계에 포함시킬 범주 범위의 마지막 범주에 해당하는 범주를 선택합니다.
2. 소계 추가...를 클릭합니다.
3. 또는 소계 정의 대화 상자에서 소계 레이블 텍스트를 수정합니다.
4. 소계만을 표시하고 소계를 정의하는 범주는 표시하지 않으려면 테이블에서 소계 범주 감추기를 선택합니다.
5. 소계를 추가하려면 계속을 클릭합니다.

총계

1. 총계 확인 상자를 클릭합니다. 총계 레이블 텍스트를 수정할 수도 있습니다.

선택한 변수가 다른 변수 내에서 중첩되는 경우 각 소계에 대한 총계가 삽입됩니다.

총계 및 소계에 대한 표시 위치

총계와 소계는 각 총계에 포함된 범주의 위 또는 아래에 표시될 수 있습니다.

- 총계 및 소계 표시에서 아래를 선택하면 총계가 각 소계의 위에 표시되고, 선택한 범주를 포함하여 그 위 (그러나 이전 소계의 아래)에 있는 모든 범주가 각 소계에 포함됩니다.
- 총계 및 소계 표시에서 위를 선택하면 총계가 각 소계의 아래에 표시되고, 선택한 범주를 포함하여 그 아래 (그러나 이전 소계의 위)에 있는 모든 범주가 각 소계에 포함됩니다.

중요: 소계를 정의하기 전에 소계의 표시 위치를 선택해야 합니다. 표시 위치 변경은 현재 선택된 소계를 제외한 모든 소계에 영향을 미치고 소계에 포함된 범주도 함께 변경됩니다.

계산된 범주

요약 통계, 총계, 소계 및/또는 상수에서 계산된 범주를 표시할 수 있습니다. 자세한 정보는 15 페이지의 『계산된 범주』 주제를 참조하십시오.

사용자 정의 총계 및 소계 요약 통계

요약 통계 대화 상자를 사용하여 테이블의 총계 및 소계 영역에 "총계" 이외의 통계를 표시할 수 있습니다. 자세한 정보는 7 페이지의 『범주형 변수의 요약 통계』 주제를 참조하십시오.

참고: 테이블의 본문에 들어 있는 사용자 정의 총계 통계를 여러 개 선택하고 통계 레이블을 감추면 총계가 테이블의 본문에서와 같은 순서로 다시 정렬되며, 레이블이 표시되지 않기 때문에 각 총계 통계가 실제로 무엇을 나타내는지 알 수 없습니다. 일반적으로 통계를 여러 개 선택하고 통계 레이블을 감추는 것은 바람직하지 않습니다.

총계, 소계 및 제외된 범주

제외된 범주의 케이스는 총계 계산에 포함되지 않습니다.

결측값, 빈 범주 및 변수값 레이블이 없는 값

결측값. 사용자 결측값 또는 결측값을 포함하도록 정의된 값(예: 남자가 임신한 경우 "해당없음"을 나타내는 코드 99)에 대한 출력을 제어합니다. 기본적으로 사용자 결측값은 제외됩니다. 테이블에 사용자 결측 범주를 포함하려면 이 옵션을 선택합니다. 변수가 여러 결측값 범주를 포함할 수 있지만 캔버스의 테이블 미리보기에는 일반 결측값 범주 하나만 표시됩니다. 테이블에는 정의된 사용자 결측 범주가 모두 포함됩니다. 시스템-결측값 (데이터 편집기에서 숫자변수의 빈 셀)은 항상 제외됩니다.

빈 범주. 빈 범주는 정의된 값 레이블이 있지만 특정 총계나 소계의 경우 해당 범주에 대한 케이스가 없는 범주입니다. 기본적으로 빈 범주는 테이블에 포함됩니다. 테이블에서 결측 범주를 제외하려면 이 옵션을 선택 취소합니다.

데이터 스캔 시 발견된 기타 값. 기본적으로 정의된 값 레이블이 없는 데이터 파일의 범주 값은 테이블에 자동으로 포함됩니다. 정의된 값 레이블이 없는 값을 테이블에서 제외하려면 이 옵션을 선택 취소합니다. 정의된 값 레이블이 있는 범주를 제외하면 정의된 값 레이블이 없는 범주도 함께 제외됩니다.

계산된 범주

테이블에는 통합된 요약 통계 결과뿐만 아니라 통합된 결과, 상수값, 소계 및 총계, 또는 이들을 조합한 항목에서 계산된 여러 범주도 표시됩니다. 이 결과를 계산된 범주 또는 사후 계산이라고 합니다. 계산된 범주는 단일 변수의 단일 범주와 마찬가지로 작동되며 다음과 같은 유사성과 차이점이 있습니다.

- 계산된 범주는 다른 범주와 같이 배치됩니다.
- 계산된 범주는 다른 범주와 동일한 통계에 사용됩니다.
- 계산된 범주는 소계, 총계 또는 유의성 검증에 영향을 주지 않습니다.
- 기본적으로, 계산된 범주의 값은 요약 통계에서 다른 범주와 동일한 형식을 사용합니다. 계산된 범주를 정의할 때 형식을 바꿀 수 있습니다.

계산된 범주는 통합된 결과의 총계를 계산하는 데 사용할 수 있으므로 소계와 유사합니다. 하지만 계산된 범주를 사용하면 소계를 사용하는 것보다 다음과 같은 이점이 있습니다.

- 계산된 범주를 다른 소계 결과에서 계산할 수 있습니다.
- 계산된 범주는 서로 중복될 수 있으므로 동일한 범주 또는 일부 동일한 범주에 사용됩니다.
- 계산된 범주는 계산된 범주 위 또는 아래에 있는 기타 모든 범주의 값을 포함할 필요가 없습니다. 즉, 계산된 범주는 전체적이 아닙니다.
- 인접하지 않은 범주의 값을 계산된 범주에 포함할 수 있습니다.

총계 및 소계와는 달리, 계산된 범주는 원래 데이터가 아닌 통합된 데이터에서 계산됩니다. 따라서 계산된 범주의 값은 총계 및 소계의 결과와 일치하지 않을 수 있습니다. 또한 계산된 범주를 정의할 때 소스 범주를 숨길 수 있는 옵션을 사용할 수 있으므로, 결과테이블에서 소계를 해석하기가 어려울 수 있습니다. 계산된 범주를 사용하는 경우 소계에 대한 사용자 정의 레이블을 지정하는 것이 좋습니다.

계산된 범주를 정의하려면

계산된 범주는 범주 및 총계 대화 상자에서 추가됩니다. 대화 상자에 액세스에 대한 자세한 내용은 12 페이지의 『범주 및 총계』 주제를 참조하십시오.

1. 범주 및 총계 대화 상자에서 범주 추가...를 클릭합니다.
2. 계산된 범주에 대한 레이블에서 계산된 범주의 레이블을 지정합니다. 범주 목록에서 범주를 끌어 범주에 대한 레이블을 포함할 수 있습니다.
3. 범주 및/또는 총계와 소계를 선택하고 계산된 범주를 정의할 연산자를 사용하여 표현식을 작성합니다. 표현식에 포함할 상수값(예: 500)을 입력할 수도 있습니다.
4. 계산된 범주만 표시하고 계산된 범주를 정의하는 범주는 표시하지 않으려면 테이블에서 표현식에 사용되는 범주 감추기를 선택합니다.
5. 표시 형식 탭을 클릭하여 계산된 범주의 표시 형식과 소수점 이하 자릿수를 변경합니다. 자세한 정보는 『계산된 범주의 표시 형식』 주제를 참조하십시오.
6. 계산된 범주를 추가하려면 계속을 클릭합니다.

계산된 범주의 표시 형식

기본적으로, 계산된 범주에서는 변수의 다른 범주와 동일한 표시 형식 및 소수점 이하 자릿수를 사용합니다. 계산된 범주 대화 상자의 표시 형식 탭에서 이러한 설정을 재지정할 수 있습니다. 표시 형식 탭에는 계산된 범주에서 사용하는 현재 요약 통계는 물론 이러한 통계의 표시 형식과 소수점 이하 자릿수가 나열됩니다.

각 요약 통계에 대해 다음을 수행할 수 있습니다.

1. 형식 드롭다운 목록에서 요약 통계에 대한 표시 형식을 선택합니다. 전체 표시 형식 목록을 보려면 11 페이지의 『요약 통계 표시 형식』 주제를 참조하십시오.
2. 선택한 요약 통계의 소수점이하자리 셀에 표시할 소수점 이하 자릿수를 입력합니다.

공유 범주가 있는 변수에 대한 테이블(Comperimeter 테이블)

설문조사에 포함된 많은 질문은 공통적인 반응 세트를 가지기도 합니다. 누적을 사용하면 이러한 관련 변수를 같은 테이블에 표시하고 공유 반응 범주를 테이블의 열에 표시할 수 있습니다.

공유 범주가 있는 다중 변수 테이블 작성

표 6. 열에 공유 반응 범주가 있는 누적변수

변수	범주 1	범주 2	범주 3
변수 1	12	34	56
변수 2	56	12	34
변수 3	34	56	12

자세한 정보는 51 페이지의 제 6 장 『공유 범주가 있는 변수 테이블』 주제를 참조하십시오.

테이블 작성기 사용자 정의

표준 대화 상자와 달리 표준 창 크기를 변경하는 것과 같은 방법으로 테이블 작성기 크기를 변경할 수 있습니다.

1. 테이블 작성기의 위쪽, 아래쪽, 측면 또는 모서리를 누른 상태로 끌어서 테이블 작성기의 크기를 줄이거나 늘립니다.

테이블 탭에서 변수 목록, 범주 목록 및 캔버스 분할창의 크기를 변경할 수도 있습니다.

2. 변수 목록과 범주 목록 사이의 수평 막대를 누른 상태로 끌어서 해당 목록을 늘이거나 줄입니다. 수평 막대를 아래로 이동하면 변수 목록은 길어지고 범주 목록은 짧아집니다. 반대로 수평 막대를 위로 이동하면 변수 목록은 짧아지고 범주 목록은 길어집니다.
3. 캔버스 분할창에서 변수 목록과 범주 목록 사이의 수직 막대를 누른 상태로 끌어서 해당 목록을 늘이거나 줄입니다. 캔버스 분할창의 크기가 나머지 공간에 맞게 자동으로 조정됩니다.

테이블 사용자 정의: 옵션 탭

데이터 셀 표시

빈 셀 및 통계를 계산할 수 없는 셀에 표시되는 내용을 제어합니다.

빈 셀 케이스가 없는 표 셀(셀 개수는 0)에서 0, 공백, 지정한 텍스트 값 등 세 가지 표시 옵션 중 하나를 선택할 수 있습니다. 텍스트 값은 최대 255자까지 사용할 수 있습니다.

계산할 수 없는 통계

통계를 계산할 수 없는 경우(예: 케이스가 없는 범주에 대한 평균)에 표시되는 텍스트입니다. 텍스트 값은 최대 255자까지 사용할 수 있습니다. 기본값은 마침표(.)입니다.

데이터 열 너비

데이터 열의 최소 및 최대 열 너비를 제어합니다. 이 설정은 행 레이블에 대한 열 너비에는 영향을 미치지 않습니다.

표모양 설정

현재 기본값 표모양의 이 Release에서는 지원되지 않음을 유의하십시오. 사용자 정의 기본값 표모양을 생성하여 새 테이블을 작성할 때 사용할 수 있으며, 표모양을 사용하여 행 레이블 열과 데이터 열 너비를 모두 제어할 수 있습니다.

사용자 정의

데이터 열 너비에 대한 기본값 표모양 설정에 우선합니다. 테이블에 대한 최소 및 최대 데이터 열 너비와 측정 단위인 점, 인치, 센티미터를 지정합니다.

척도변수의 결측값

척도변수가 두 개 이상 있는 테이블에서 척도변수 통계에 대한 결측데이터 처리 방법을 제어합니다.

데이터 사용의 최대화(변수별 삭제)

각 척도변수에 대한 유효값이 있는 모든 케이스가 해당 척도변수에 대한 요약 통계에 포함됩니다.

전체 척도변수에 대해 일관된 케이스 기준 사용(목록별 삭제)

테이블의 척도변수에 대한 결측값이 있는 케이스는 테이블의 모든 척도변수에 대한 요약 통계에서 제외됩니다.

유효 기준

빈도 가중치가 아니라 조정 가중치를 나타내는 변수가 있는 경우 해당 변수를 유효 기준 가중치 변수로 사용할 수 있습니다. 유효 기준 또는 유효 샘플 크기 가중치의 개념은 복합 샘플의 데이터 분석 방

법을 기반으로 합니다. 유효 기준 가중치를 사용하면 조정 가중치를 사용하여 단순 랜덤 샘플링 설계의 임시 데이터 조정과 관련된 분석에서 통계적 추론의 추정 처리가 가능합니다.

- 유효 기준 가중치는 가중치가 부여된 요약 통계 값과 열 평균 및 열 비율 유의성 검정에 영향을 미칩니다.
- 데이터 세트에 대해 가중치가 적용되는 경우 데이터 세트 가중치 변수가 무시되고 유효 기준 가중치 변수로 가중치가 부여됩니다.
- 유효 기준 가중치 변수는 고유해야 합니다.
- 가중치 값이 음의 값이거나 0인 케이스 또는 가중치 값이 누락된 케이스는 모든 결과에서 제외됩니다.

다중 범주형 변수군의 중복 반응수 계산

중복 반응은 다중 범주 세트에서 두 개 이상의 변수에 동일한 반응입니다. 기본적으로 중복 반응수는 계산되지 않습니다.

작은 수 숨기기

지정된 정수보다 작은 수를 숨기도록 선택할 수 있습니다.

- 숨겨진 값은 <N으로 표시되며 여기서 N은 지정된 정수입니다.
- 지정된 정수는 2보다 크거나 같아야 합니다.
- 데이터 세트에 대해 가중치가 적용되거나 유효 기준 가중치 변수가 지정된 경우에는 가중치가 부여된 값이 사용됩니다.

가중치 및 반올림

- 데이터 > 케이스 가중치를 사용하여 케이스에 가중치를 부여하는 경우, 정수가 아닌 가중치는 유의성 검정, 신뢰구간 및 표준 오류의 셀 또는 범주 수준에서 반올림됩니다.
- 유효 기준 가중치 변수 사용을 선택하는 경우, 선택한 가중치 변수에서 정수가 아닌 가중치 값은 반올림되지 않습니다.
- 둘 모두 지정되는 경우, 유효 기준 가중치 변수가 사용되고 정수가 아닌 가중치는 반올림되지 않습니다.

테이블 사용자 정의: 제목 탭

제목 탭은 제목, 캡션 및 코너 레이블의 출력을 제어합니다.

제목. 테이블 위에 표시되는 텍스트입니다.

캡션. 테이블 아래 꼬리말 위에 표시되는 텍스트입니다.

코너. 테이블 왼쪽 상단에 표시되는 텍스트입니다. 코너 텍스트는 테이블에 행 변수가 있는 경우와 피벗 테이블 행 차원 레이블 특성이 중첩으로 설정된 경우에만 표시되며 기본값 표모양 설정이 아닙니다.

테이블 제목, 캡션 또는 코너 레이블에 다음과 같은 자동으로 생성된 값을 포함할 수 있습니다.

날짜. 현재 Windows 국가별 옵션 설정에 지정된 형식으로 표시되는 현재 년, 월, 일입니다.

시간. 현재 Windows 국가별 옵션 설정에 지정된 형식으로 표시되는 현재 시, 분, 초입니다.

표 표현식. 테이블에 사용되는 변수 및 변수가 테이블에 사용되는 방법입니다. 변수에 정의된 변수 레이블이 있는 경우 해당 레이블이 표시됩니다. 생성된 테이블에서 다음 기호는 변수가 테이블에 사용되는 방법을 나타냅니다.

- +는 변수 누적을 나타냅니다.
- >은(는) 중첩을 나타냅니다.
- **BY**는 교차 분석표 또는 레이어를 나타냅니다.

테이블 사용자 정의: 검정 통계량 탭

검정 통계량 탭에서는 사용자 테이블에 대한 유의성 검정을 제공합니다.

범주 레이블을 해당 기본값 테이블 차원에서 이동하거나 계산된 범주에 대해 이동한 테이블에 대해서는 이러한 검정을 사용할 수 없습니다.

열 평균 및 열 비율 검정

열 평균 검정은 척도변수에 사용할 수 있습니다. 열 비율 검정은 범주형 변수에 사용할 수 있습니다.

열 평균 비교

열 평균의 등식에 대한 대응별 검정입니다. 테이블의 열에는 범주형 변수가 포함되고 가장 안쪽 레벨의 행으로 척도변수가 포함되어야 합니다. 테이블에는 평균이 요약 통계로 포함되어야 합니다.

일반 범주형 변수의 경우, 모든 범주에서 또는 비교되는 범주에서만 분산을 추정할 수 있습니다. 다중 반응 변수의 경우, 평균 검정의 분산은 항상 비교되는 범주만을 기반으로 합니다.

열 비율 비교

열 비율의 등식에 대한 대응별 검정입니다. 테이블의 열 및 행 모두에 하나 이상의 범주형 변수가 포함되어야 합니다. 테이블에는 개수 또는 열 퍼센트가 포함되어야 합니다.

유의차 식별

열 평균 및 열 비율 검정의 경우, 별도의 테이블 또는 기본 테이블에 유의한 결과를 표시할 수 있습니다.

별도의 테이블에서

유의성 검증 결과가 별도의 테이블에 표시됩니다. 두 값에 상당한 차이가 있는 경우, 큰 값에 해당하는 셀은 작은 값이 포함된 열을 식별하는 키를 표시합니다.

유의성 값 표시

유의성 값은 셀에서 각 키 값 뒤의 소괄호 안에 표시됩니다. 이 옵션은 유의한 결과가 별도의 테이블에 표시될 때만 사용할 수 있습니다.

기본 테이블에서

유의성 검정 결과가 기본 테이블에 표시됩니다. 테이블의 각 열 범주는 영문자 키로 식별됩니다. 유의한 쌍마다 열 평균 또는 열 비율이 더 작은 범주의 키가 열 평균 또는 열 비율이 더 큰 범주에 표시됩니다.

- 피벗 테이블의 열 레이블 셀에서 키 위로 커서를 가져가면 해당 유의성 키를 가진 테이블의 모든 셀이 강조표시됩니다. 열 차원에 여러 변수가 있는 테이블의 경우, 해당 하위 테이블의 셀만 강조표시됩니다.
- 동일한 유의성 키를 가진 테이블(또는 하위 테이블)의 모든 셀을 선택하려면 열 레이블 셀을 클릭하고 선택 > 이 유의성 키가 있는 모든 셀 선택을 선택합니다.

APA 스타일 아래첨자 사용

아래첨자를 사용하는 APA 스타일 형식화를 사용하여 유의차를 식별합니다. 두 값이 현저히 다른 경우 두 값은 서로 다른 아래첨자를 표시합니다. 아래첨자는 꼬리말이 아닙니다. 이 옵션이 적용되고 있는 경우, 현재 표모양에 정의된 꼬리말 스타일이 재지정되고 꼬리말이 윗첨자 숫자로 표시됩니다. 동일한 유의성 키를 가진 동일한 행의 모든 셀을 선택하려면 유의성 키가 있는 셀을 마우스 오른쪽 단추로 클릭하고 유의성이 비슷한 셀 선택을 선택합니다.

유의 수준

열 평균 및 열 비율 검정의 유의수준입니다.

- 이 값은 0보다 크고 1보다 작아야 합니다.
- 두 개의 유의수준을 지정하는 경우, 더 작은 수준보다 작거나 같은 유의값을 식별하기 위해 대문자가 사용됩니다. 소문자는 더 큰 수준보다 작거나 같은 유의값을 식별하는 데 사용됩니다.
- APA 스타일 아래첨자 사용을 선택하면 두 번째 값은 무시됩니다.

다중비교를 위해 P-값 조정

Bonferroni 보정은 FWER(family-wise error rate)을 조정합니다. **Benjamini-Hochberg** 방법은 FDR(false discovery rate) 조정입니다. 이 방법은 Bonferroni 보정보다 덜 보수적(conservative)입니다.

독립성 검정(카이제곱)

행과 열 모두에 하나 이상의 범주형 변수가 있는 테이블에 대한 카이제곱 독립성 검정입니다.

소계 범주 대신 소계 사용

각 소계가 유의성 검정을 위한 범주를 대체합니다. 선택하지 않으면 소계 범주가 감춰지는 소계만이 검정을 위한 범주를 대체합니다.

검정에 다중 반응 변수 포함

유의성 검정에 다중 반응 세트의 범주가 포함됩니다. 그렇지 않은 경우 다중 반응 세트가 유의성 검정에 포함되지 않습니다.

제 2 장 범주형 변수에 대한 단순 통계표












범주형 변수에 대한 단순 통계표

작성하는 대부분의 테이블에는 범주형 변수가 하나 이상 포함됩니다. 범주형 변수란 고유한 값이나 범주 수가 제한된 변수(예: 성별 또는 종교)입니다. 범주형 변수는 명목이거나 순서일 수 있습니다.

- **명목(Nominal)**. 변수의 값이 고유한 순위가 없는 범주를 나타내는 경우 해당 변수는 명목으로 취급될 수 있습니다. 예를 들어, 직원이 근무하는 회사의 부서가 있습니다. 종교, 우편번호 또는 종교 단체 등이 명목 변수에 해당합니다.
- **순서(Ordinal)**. 변수의 값이 고유한 순위가 있는 범주를 나타내는 경우 해당 변수는 순서로 취급될 수 있습니다. 예를 들어, 매우 불만족에서 매우 만족에 이르는 서비스 만족도 수준이 있습니다. 순서변수의 예로는 만족도나 신뢰도를 나타내는 태도 스코어 및 선호도 등급 스코어가 있습니다.

변수 목록의 각 변수 옆에 있는 아이콘은 변수 유형을 식별합니다.

표 7. 측정 수준 아이콘

	숫자	문자열	날짜	시간
척도(연속형)		해당 없음		
순서				
명목				

테이블 사용자 정의는 정의된 값 레이블이 있는 범주형 변수를 사용할 때 가장 효율적입니다. 자세한 정보는 1 페이지의 『테이블 작성』 주제를 참조하십시오.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 설정합니다.

단일 범주형 변수

단일 범주형 변수에 대한 통계표는 작성할 수 있는 가장 단순한 표이지만 이 통계표만으로도 충분한 경우가 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

캔버스 분할창에 테이블의 미리보기가 표시됩니다. 미리보기에는 실제 데이터 값이 표시되지 않고 데이터가 표시될 자리표시자만 표시됩니다.

3. 확인을 클릭하여 테이블을 작성합니다.

표가 뷰어 창에 표시됩니다.

		Count
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

그림 2. 행으로 표시한 단일 범주형 변수

단순 통계표에서 열 머리말인 개수는 실제로 필요하지 않으며 열 머리말 없이 테이블을 작성할 수 있습니다.

4. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
5. 요약 통계의 위치에서 숨김을 선택합니다.
6. 확인을 클릭하여 테이블을 작성합니다.

Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

그림 3. 요약 통계 열 레이블이 없는 단일 범주형 변수

퍼센트

개수 이외에 퍼센트도 표시할 수 있습니다. 단일 범주형 변수에 대한 단순 통계표에서 변수가 행으로 표시되는 경우 열 퍼센트를 표시할 수 있습니다. 반대로 변수가 열로 표시되는 경우 행 퍼센트를 표시할 수 있습니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).

2. 요약 통계의 위치에서 숨김을 선택 취소합니다. 이 테이블에는 열이 두 개이므로 각 열이 무엇을 나타내는 지 알 수 있도록 열 레이블을 표시할 수 있습니다.
3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
4. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
5. 출력 목록의 레이블 셀에서 기본값 레이블을 삭제하고 퍼센트를 입력합니다.
6. 선택한 항목에 적용 을 클릭한 다음 테이블 작성기에서 확인을 클릭하여 테이블을 작성합니다.

		Count	Percent
Age category	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%

그림 4. 개수 및 열 퍼센트

총계

총계는 테이블 사용자 정의에 자동으로 포함되지 않지만 테이블에 총계를 추가하는 것은 간단합니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
3. 범주 및 총계 대화 상자에서 총계 확인란을 선택합니다.
4. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Percent
Age category	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%
Total		2828	100.0%

그림 5. 개수, 열 퍼센트 및 총계

자세한 정보는 37 페이지의 제 4 장 『범주형 변수에 대한 총계 및 소계』 주제를 참조하십시오.

교차 분석표

교차 분석표는 두 범주형 변수 간의 관계를 검토하는 기본적인 기법입니다. 예를 들어, 행 변수로 연령범주를 사용하고 열 변수로 성별을 사용하면 각 연령 범주에 대한 여자 및 남자의 수를 표시하는 2차원 교차 분석표를 만들 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).

2. 재설정을 클릭하여 테이블 작성기의 이전 선택을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 성별을 캔버스 분할창의 열 영역으로 끌어다 놓습니다. 이 변수를 찾으려면 변수 목록을 아래로 스크롤해야 합니다.
5. 확인을 클릭하여 테이블을 작성합니다.

		Gender	
		Male	Female
		Count	Count
Age category	Less than 25	108	134
	25 to 34	276	351
	35 to 44	309	370
	45 to 54	221	260
	55 to 64	136	184
	65 or older	178	301

그림 6. 연령범주 및 성별에 대한 교차 분석표

교차 분석표의 퍼센트

2차원 교차 분석표의 열 및 행 퍼센트는 유용한 정보를 제공합니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 성별을 마우스 오른쪽 단추로 클릭합니다.

팝업 메뉴에서 요약 통계를 사용할 수 없음을 알 수 있습니다. 통계 소스 차원에서 맨 인쪽에 있는 변수에 대한 요약 통계만 선택할 수 있기 때문입니다. 범주형 변수에 대한 기본값 통계 소스 차원(행 또는 열)은 캔버스 분할창으로 변수를 끌어다 놓는 순서를 기준으로 합니다. 이 예제에서는 먼저 연령범주를 행 차원으로 끌어다 놓았고 행 차원에는 다른 변수가 없으므로 연령범주가 통계 소스 변수가 됩니다. 통계 소스 차원을 변경할 수 있지만 이 예제에서는 그럴 필요가 없습니다. 자세한 정보는 6 페이지의 『요약 통계』 주제를 참조하십시오.

3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
4. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
5. 통계 목록에서 행 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
6. 선택한 항목에 적용 을 클릭한 다음 테이블 작성기에서 확인을 클릭하여 테이블을 작성합니다.

		Gender					
		Male			Female		
		Count	Column N %	Row N %	Count	Column N %	Row N %
Age category	Less than 25	108	8.8%	44.6%	134	8.4%	55.4%
	25 to 34	276	22.5%	44.0%	351	21.9%	56.0%
	35 to 44	309	25.2%	45.5%	370	23.1%	54.5%
	45 to 54	221	18.0%	45.9%	260	16.3%	54.1%
	55 to 64	136	11.1%	42.5%	184	11.5%	57.5%
	65 or older	178	14.5%	37.2%	301	18.8%	62.8%

그림 7. 행 및 열 퍼센트가 있는 교차 분석표

표시 형식 제어

요약 통계에 출력된 소수점 이하 자릿수를 비롯해 출력 형식을 제어할 수 있습니다. 예를 들어, 퍼센트는 기본적으로 소수점 이하 한 자리와 퍼센트 기호로 출력되지만 셀 값이 퍼센트 기호 없이 소수점 이하 두 자리로 표시되도록 할 수도 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
3. 선택한 두 개의 퍼센트 요약 통계(열 N % 및 행 N %)에 대해 형식 드롭다운 목록에서 **nnnn.n**을 선택하고 소수점이하자리 셀에 2를 입력합니다.
4. 확인을 클릭하여 테이블을 작성합니다.

		Gender					
		Male			Female		
		Count	Column N %	Row N %	Count	Column N %	Row N %
Age category	Less than 25	108	8.79	44.63	134	8.38	55.37
	25 to 34	276	22.48	44.02	351	21.94	55.98
	35 to 44	309	25.16	45.51	370	23.13	54.49
	45 to 54	221	18.00	45.95	260	16.25	54.05
	55 to 64	136	11.07	42.50	184	11.50	57.50
	65 or older	178	14.50	37.16	301	18.81	62.84

그림 8. 행 및 열 퍼센트에 대해 형식 지정된 셀 출력

주변 합계

교차 분석표에서 주변 합계 즉, 각 행과 열의 합계를 표시하는 경우는 아주 흔합니다. 주변 합계는 테이블 사용자 정의에 기본적으로 포함되지 않으므로 테이블에 명시적으로 추가해야 합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 재설정을 클릭하여 테이블 작성기의 이전 선택을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 성별을 캔버스 분할창의 열 영역으로 끌어다 놓습니다. 이 변수를 찾으려면 변수 목록을 아래로 스크롤해야 합니다.

5. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
6. 범주 및 총계 대화 상자에서 총계 확인란을 선택한 다음 적용을 클릭합니다.
7. 캔버스 분할창에서 성별을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
8. 범주 및 총계 대화 상자에서 총계 확인란을 선택한 다음 적용을 클릭합니다.
9. 요약 통계의 위치에서 숨김을 선택합니다. 개수만 출력할 것이므로 테이블의 데이터 셀에 출력된 "통계"는 식별하지 않아도 됩니다.
10. 확인을 클릭하여 테이블을 작성합니다.

		Gender		
		Male	Female	Total
Age category	Less than 25	108	134	242
	25 to 34	276	351	627
	35 to 44	309	370	679
	45 to 54	221	260	481
	55 to 64	136	184	320
	65 or older	178	301	479
	Total	1228	1600	2828

그림 9. 주변 합계가 있는 교차 분석표

범주 정렬 및 제외

기본적으로 범주에서는 범주 값 레이블이 나타내는 데이터 값을 오름차순으로 표시합니다. 예를 들어, 연령범주에 대해 25 미만, 25 ~ 34, 35 ~ 44 등의 값 레이블이 표시되지만 실제 기본 데이터 값은 1, 2, 3 등이며 이 기본 데이터 값에 따라 범주의 출력 순서가 결정됩니다.

범주 순서를 쉽게 변경할 수 있으며 테이블에 표시하지 않을 범주를 제외할 수도 있습니다.

범주 정렬

다음 값을 기준으로 오름차순 또는 내림차순으로 범주를 정렬하거나 재정렬할 수 있습니다.

- 데이터 값
- 값 레이블.
- 셀 개수
- 요약 통계. 정렬에 사용할 수 있는 요약 통계는 테이블에 표시할 요약 통계에 따라 다릅니다.
 1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
 2. 연령범주가 캔버스 분할창의 행 영역에 표시되어 있지 않으면 여기로 끌어다 놓습니다.
 3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.

데이터 값과 해당 값 레이블이 현재의 정렬 순서인 데이터 값의 오름차순으로 표시됩니다.

4. 범주 정렬의 순서 드롭다운 목록에서 내림차순을 선택합니다.

정렬 순서가 뒤바뀝니다.

5. 기준 드롭다운 목록에서 레이블을 선택합니다.

이제 범주가 값 레이블의 내림차순 문자순으로 정렬됩니다.

25 미만이라는 레이블이 붙은 범주가 목록 맨 위에 표시됩니다. 문자순 정렬에서는 문자가 숫자 다음에 옵니다. 이 레이블만 문자로 시작하고 목록이 내림차순(반대로) 정렬되기 때문에 이 범주가 목록에서 맨 위에 정렬됩니다.

특정 범주를 목록의 다른 위치에 표시하려면 범주를 옮기면 됩니다.

6. 레이블 목록에서 25 미만으로 설명된 범주를 클릭합니다.

7. 목록 오른쪽의 아래쪽 화살표를 클릭합니다. 목록의 범주가 한 행씩 아래로 이동합니다.

8. 목록 맨 아래의 범주에 도달할 때까지 아래쪽 화살표를 계속 클릭합니다.

범주 제외

테이블에서 표시하지 않을 범주가 있는 경우 제외할 수 있습니다.

1. 레이블 목록에서 25 미만으로 설명된 범주를 클릭합니다.

2. 제외 목록 왼쪽의 화살표 키를 클릭합니다.

3. 레이블 목록에서 65 이상으로 설명된 범주를 클릭합니다.

4. 제외 목록 왼쪽의 화살표 키를 다시 클릭합니다.

두 범주가 출력 목록에서 제외 목록으로 이동됩니다. 범주를 출력 목록으로 쉽게 이동시킬 수 있습니다.

5. 테이블 작성기에서 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Gender		
		Male	Female	Total
Age category	55 to 64	136	184	320
	45 to 54	221	260	481
	35 to 44	309	370	679
	25 to 34	276	351	627
	Total	942	1165	2107

그림 10. 일부 범주를 제외하고 값 레이블의 내림차순으로 정렬한 테이블

두 범주를 제외하기 전보다 총계가 더 적어졌음을 알 수 있습니다. 이는 테이블에 포함된 범주에 따라 총계를 계산하기 때문입니다. 제외된 범주는 총계 계산에서 제외됩니다. 자세한 정보는 37 페이지의 제 4 장 『범주형 변수에 대한 총계 및 소계』 주제를 참조하십시오.

제 3 장 범주형 변수를 사용한 누적, 중첩 및 레이어

누적, 중첩 및 레이어는 모두 다중 변수를 같은 테이블에 표시하기 위한 방법입니다. 이 장 은 척도변수에서도 사용할 수 있지만 이 장에서는 범주형 변수에서 사용하는 경우에 중점을 두어 설명합니다.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 설정합니다.

범주형 변수 누적

누적은 개별 테이블을 취하여 동일한 출력 화면에 붙여넣는 것입니다. 예를 들어, 같은 테이블에서 **성별**과 **연령범주**에 대한 정보를 별도의 섹션으로 표시할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록의 **성별**을 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

3. 변수 목록에서 **연령범주**를 **성별** 아래의 행 영역으로 끌어다 놓습니다.

두 변수가 행 차원에 누적됩니다.

4. 확인을 클릭하여 테이블을 작성합니다.

		Count
Gender	Male	1232
	Female	1600
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

그림 11. 행으로 누적된 범주형 변수에 대한 테이블

같은 방법으로 변수를 열로 누적할 수도 있습니다.

교차 분석표에서의 누적

누적표에서는 차원마다 다른 변수를 포함할 수 있습니다. 예를 들어, 두 변수를 행으로 누적하고 세 번째 변수를 열 차원에 표시한 교차 분석표를 작성할 수 있습니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 연령범주와 성별가 행으로 누적되어 있지 않으면 위의 방법에 따라 누적합니다.
3. 변수 목록의 인터넷으로부터 뉴스얻음을 캔버스 분할창의 열 영역으로 끌어다 놓습니다.
4. 확인을 클릭하여 테이블을 작성합니다.

		Get news from internet	
		No	Yes
		Count	Count
Gender	Male	873	359
	Female	1092	508
Age category	Less than 25	146	96
	25 to 34	368	259
	35 to 44	435	244
	45 to 54	346	135
	55 to 64	252	68
	65 or older	416	63

그림 12. 누적된 행 변수 두 개와 열 변수 하나로 구성된 교차 분석표

참고: 변수 목록에 ... (으)로부터 뉴스 얻음이라는 레이블을 가진 변수가 여러 개가 있을 수 있는데 이 레이블이 너무 길어서 레이블 전체가 변수 목록에 표시되지 않을 수 있습니다. 이런 경우 변수 목록에서 개별 변수를 정확하게 구분하기가 어렵습니다. 전체 변수 레이블을 보는 방법에는 다음 두 가지가 있습니다.

- 마우스 포인터를 목록의 변수 위에 놓으면 전체 레이블이 팝업 도구 팁에 표시됩니다.
- 캔버스 분할창에서 변수와 범주 목록을 구분하는 수직 막대를 누른 상태로 끌어서 목록을 넓힙니다.

중첩 범주형 변수

두 범주형 변수 간의 관계를 표시할 수 있다는 점에서는 중첩은 교차 분석표와 비슷하지만, 교차 분석표와 달리 중첩에서는 한 변수가 같은 차원의 다른 변수 내에 중첩됩니다. 예를 들어, 행 차원의 연령범주 내에 성별을 중첩하여 각 연령 범주에 대한 여자 및 남자의 수를 표시할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 재설정을 클릭하여 테이블 작성기의 이전 선택을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 성별을 행 영역의 연령범주 오른쪽으로 끌어다 놓습니다.

이제 캔버스 분할창의 미리보기에는 단일 개수 열이 포함된 중첩 테이블이 표시되며 각 셀에는 연령범주에 따라 여자 또는 남자 수가 포함됩니다.

여기서 성별이라는 변수 레이블이 각 연령범주에 대해 반복해서 표시되는 것을 알 수 있습니다. 범주 수가 가장 적은 변수를 맨 바깥쪽 중첩 수준에 넣으면 이러한 반복을 최소화할 수 있습니다.

5. 캔버스 분할창에서 변수 레이블 성별을 클릭합니다.
6. 이 변수를 맨 왼쪽 행 영역으로 끌어다 놓습니다.

이제 성별이 여섯 번 반복되는 대신 연령범주가 두 번 반복됩니다. 이렇게 하면 기본적으로 동일한 결과를 표시하지만 산만하지 않은 테이블을 작성할 수 있습니다.

7. 확인을 클릭하여 테이블을 작성합니다.

				Count
Gender	Male	Age category	Less than 25	108
			25 to 34	276
			35 to 44	309
			45 to 54	221
			55 to 64	136
	Female	Age category	65 or older	178
			Less than 25	134
			25 to 34	351
			35 to 44	370
			45 to 54	260
			55 to 64	184
			65 or older	301

그림 13. 성별 내에 중첩된 연령범주에 대한 테이블

참고: 테이블 사용자 정의는 레이어 분할 파일 처리를 사용하지 않습니다. 레이어 분할 파일과 같은 결과를 얻으려면 분할 파일 변수를 테이블의 가장 바깥쪽에 있는 중첩 레이어에 두십시오.

변수 레이블 출력안음

중첩된 테이블의 중복된 변수 레이블에 대한 다른 해법으로 단순히 변수 이름 또는 레이블을 표시하지 않는 방법이 있습니다. 성별 및 연령범주에 대한 값 레이블의 경우 변수 레이블이 없어도 충분히 알아 볼 수 있으므로 두 변수에 대한 레이블을 제거할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 변수 레이블 표시를 선택 취소합니다.
3. 성별에도 같은 작업을 수행합니다.

변수 레이블이 테이블 미리보기에 여전히 표시되지만 테이블에는 포함되지 않습니다.

4. 확인을 클릭하여 테이블을 작성합니다.

		Count
Male	Less than 25	108
	25 to 34	276
	35 to 44	309
	45 to 54	221
	55 to 64	136
	65 or older	178
Female	Less than 25	134
	25 to 34	351
	35 to 44	370
	45 to 54	260
	55 to 64	184
	65 or older	301

그림 14. 변수 레이블이 없는 중첩된 테이블

변수 레이블을 테이블 내용에 여러 번 표시하지 않고 테이블의 다른 곳에 포함시키려는 경우 테이블 제목이나 코너 레이블에 포함시킬 수 있습니다.

5. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
6. 제목 탭을 클릭합니다.
7. 제목 입력란에서 아무 위치나 클릭합니다.
8. 표 표현식을 클릭합니다. 텍스트 &[표 표현식]이 제목 입력란에 표시됩니다. 이렇게 하면 테이블에서 사용된 변수 레이블이 포함된 테이블 제목이 생성됩니다.
9. 확인을 클릭하여 테이블을 작성합니다.

Gender > Age category

		Count
Male	Less than 25	108
	25 to 34	276
	35 to 44	309
	45 to 54	221
	55 to 64	136
	65 or older	178
Female	Less than 25	134
	25 to 34	351
	35 to 44	370
	45 to 54	260
	55 to 64	184
	65 or older	301

그림 15. 테이블 제목으로 표시된 변수 레이블

제목에서 기호(>)는 연령범주가 성별 내에 중첩되어 있다는 것을 나타냅니다.

중첩 교차 분석표

중첩 테이블에서는 차원마다 다른 변수를 포함할 수 있습니다. 예를 들어, 행에서 연령범주를 성별 내에 중첩한 다음 이 중첩 행과 열 차원의 세 번째 변수로 교차 분석표를 작성할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 연령범주가 행의 성별 내에 이미 중첩되어 있지 않으면 이전 예제의 방법을 따라 중첩합니다.
3. 변수 목록의 인터넷으로부터 뉴스언음을 캔버스 분할창의 열 영역으로 끌어다 놓습니다.

표가 너무 커서 캔버스 분할창에 전체가 나타나지 않을 수도 있습니다. 테이블 미리보기에서 많은 부분을 보려면 상하 또는 좌우로 스크롤하거나 다음과 같이 하십시오.

- 테이블 작성기에서 **최소**를 클릭하여 최소 보기를 표시합니다. 여기서는 테이블에 포함된 범주 또는 요약 통계에 대한 정보 없이 변수 레이블만 표시합니다.
- 테이블 작성기의 옆면이나 코너를 누른 채 끌어서 테이블 작성기의 크기를 늘립니다.

4. **확인**을 클릭하여 테이블을 작성합니다.

				Get news from internet	
				No	Yes
				Count	Count
Gender	Male	Age category	Less than 25	59	49
			25 to 34	159	117
			35 to 44	217	92
			45 to 54	169	52
			55 to 64	112	24
			65 or older	155	23
			Female	Age category	Less than 25
			25 to 34	209	142
			35 to 44	218	152
			45 to 54	177	83
			55 to 64	140	44
			65 or older	261	40

그림 16. 중첩 교차 분석표

행과 열 바꾸기

시간을 들여 복잡한 테이블을 완벽하게 만들었는데 모든 행 변수를 열로 이동하거나 열 변수를 행으로 이동하는 등 방향을 완전히 바꾸려는 경우에는 다음과 같이 하십시오. 예를 들어, 연령범주와 성별을 행에 중첩하여 교차 분석표를 작성했지만 이 두 가지 인구 통계학적 변수를 열에 중첩하려고 한다고 가정합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 아무 위치나 마우스 오른쪽 단추로 클릭한 다음 팝업 메뉴에서 **행 및 열 변수 바꾸기**를 선택합니다.

행 변수와 열 변수가 바뀌었습니다.

테이블을 작성하기 전에 산만하지 않도록 몇 가지 수정합니다.

3. 요약 통계 열 레이블을 출력하지 않으려면 숨김을 선택합니다.
4. 캔버스 분할창에서 성별을 마우스 오른쪽 단추로 클릭하고 변수 레이블 표시를 선택 취소합니다.
5. 이제 **확인**을 클릭하여 테이블을 작성합니다.

		Male						Female					
		Age category						Age category					
		Less than 25	25 to 34	35 to 44	45 to 54	55 to 64	65 or older	Less than 25	25 to 34	35 to 44	45 to 54	55 to 64	65 or older
Get news from internet	No	59	159	217	169	112	155	87	209	218	177	140	261
	Yes	49	117	92	52	24	23	47	142	152	83	44	40

그림 17. 인구 통계학적 변수가 열에 중첩된 교차 분석표

레이어

레이어를 사용하면 테이블에 차원 깊이를 추가하여 3차원 "큐브"를 만들 수 있습니다. 레이어는 중첩이나 누적과 유사하지만 레이어 범주는 한 번에 하나만 표시할 수 있다는 점에서 다릅니다. 예를 들어, 행 변수로 연령범주를 사용하고 레이어 변수로 성별을 사용하면 여자와 남자에 대한 정보가 다른 레이어에 표시되는 테이블을 작성할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 재설정을 클릭하여 테이블 작성기의 이전 선택을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 테이블 작성기 위쪽에 있는 레이어를 클릭하여 레이어 목록을 표시합니다.
5. 변수 목록에서 성별을 레이어 목록으로 끌어다 놓습니다.

여기서 레이어 변수를 추가해도 캔버스 분할창에 표시된 미리보기에는 아무 영향이 없음을 알 수 있습니다. 레이어 변수가 통계 소스 변수이고 요약 통계를 변경하는 경우 외에는 레이어 변수는 캔버스 분할창의 미리보기에 영향을 주지 않습니다.

6. 확인을 클릭하여 테이블을 작성합니다.

		Count
Age category	Less than 25	108
	25 to 34	276
	35 to 44	309
	45 to 54	221
	55 to 64	136
	65 or older	178

그림 18. 단순 레이어 테이블

이 테이블은 단일 범주형 변수에 대한 단순 통계표와 크게 차이가 없습니다. 테이블 위쪽에 성별 남자라는 레이블이 있다는 점만 다릅니다.

7. 뷰어 창에서 테이블을 두 번 클릭하여 활성화합니다.
8. 이제 성별 남자라는 레이블이 실제로 드롭다운 목록의 선택 항목으로 되어 있음을 확인할 수 있습니다.
9. 드롭다운 목록에서 아래쪽 화살표를 클릭하여 레이어 전체 목록을 표시합니다.

이 테이블에서는 목록에 다른 선택 항목이 하나뿐입니다.

10. 드롭다운 목록에서 **성별** 여지를 선택합니다.

		Count
Age category	Less than 25	134
	25 to 34	351
	35 to 44	370
	45 to 54	260
	55 to 64	184
	65 or older	301

그림 19. 다른 레이어가 표시된 단순 레이어 테이블

두 개의 누적 범주형 레이어 변수

레이어에 둘 이상의 범주형 변수가 있는 경우 레이어 변수를 누적 또는 중첩할 수 있습니다. 기본적으로 레이어 변수는 누적됩니다. (참고: 척도 레이어 변수가 있는 경우 레이어 변수를 누적으로만 사용할 수 있습니다.)

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 행에 연령범주가 없고 레이어에 성별이 없는 경우 이전 예제의 방법을 따라 레이어 테이블을 작성합니다.
3. 변수 목록에서 최종학력을 레이어 목록의 성별 아래쪽으로 끌어다 놓습니다.

레이어 목록 아래에 있는 레이어 결과에서 두 개의 단일 선택 단추가 활성화됩니다. 기본값 선택은 각 범주를 레이어로 표시이며 이는 누적과 동일합니다.

4. 확인을 클릭하여 테이블을 작성합니다.
5. 뷰어 창에서 테이블을 두 번 클릭하여 활성화합니다.
6. 드롭다운 목록에서 아래쪽 화살표를 클릭하여 레이어 전체 목록을 표시합니다.

테이블에는 7 개의 레이어가 있습니다. 두 개의 레이어는 두 개의 성별 범주를 위한 것이고 다섯 개의 레이어는 다섯 개의 최종학력 범주를 위한 것입니다. 누적 레이어의 경우 전체 레이어 수는 레이어 변수에 대해 요청한 총계 또는 소계 범주를 포함하여 레이어 변수에 대한 범주 수의 합입니다.

두 개의 중첩된 범주형 레이어 변수

범주형 레이어 변수를 중첩하면 레이어 변수 범주마다 다른 레이어를 만듭니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 아직 만들지 않은 경우 이전 예제의 방법을 따라 누적 레이어 테이블을 작성합니다.
3. 레이어 결과에서 각 범주 조합을 레이어로 표시를 선택합니다. 이는 중첩과 동일합니다.
4. 확인을 클릭하여 테이블을 작성합니다.
5. 뷰어 창에서 테이블을 두 번 클릭하여 활성화합니다.
6. 드롭다운 목록에서 아래쪽 화살표를 클릭하여 레이어 전체 목록을 표시합니다.

테이블에는 성별 및 최종학력의 조합마다 하나씩 총 10개의 레이어가 있습니다. 전체 목록을 보려면 스크롤해야 합니다. 중첩 레이어의 경우 전체 레이어 수는 각 레이어 변수에 대한 범주 수의 결과입니다. 이 예제의 경우 $5 \times 2 = 10$ 입니다.

레이어 테이블 인쇄

기본적으로 현재 보이는 레이어만 인쇄됩니다. 테이블의 모든 레이어를 인쇄하려면 다음과 같이 하십시오.

1. 뷰어 창에서 테이블을 두 번 클릭하여 활성화합니다.
2. 뷰어 창 메뉴에서 다음을 선택합니다:

형식 > 테이블 특성...

3. 인쇄 탭을 클릭합니다.
4. 모든 레이어 인쇄를 선택합니다.

기본값 표모양을 포함하여 이 설정을 표모양의 일부로 저장할 수도 있습니다.

제 4 장 범주형 변수에 대한 총계 및 소개

테이블 사용자 정의에 총계와 소개를 모두 포함할 수 있습니다. 총계와 소개는 행, 열, 레이어 등 모든 차원의 모든 중첩 수준에서 범주형 변수에 적용될 수 있습니다.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 설정합니다.

단일 변수의 단순 총계

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
4. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
5. 출력 목록의 레이블 셀에서 기본값 레이블을 삭제하고 퍼센트를 입력합니다.
6. 선택한 항목에 적용을 클릭합니다.
7. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
8. 범주 및 총계 대화 상자에서 총계 확인란을 선택합니다.
9. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Percent
Age category	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%
	Total	2828	100.0%

그림 20. 단일 범주형 변수의 단순 총계

표시된 내용대로 총계 생성

테이블에 표시된 범주에 따라 총계가 계산됩니다. 테이블에서 일부 범주를 제외하도록 선택하면 해당 범주의 케이스가 총계 계산에 포함되지 않습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
3. 레이블 목록에서 25 미만으로 설명된 범주를 클릭합니다.
4. 제외 목록 왼쪽의 화살표 키를 클릭합니다.
5. 레이블 목록에서 65 이상으로 설명된 범주를 클릭합니다.
6. 제외 목록 왼쪽의 화살표 키를 다시 클릭합니다.

두 범주가 출력 목록에서 제외 목록으로 이동됩니다.

7. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Percent
Age category	25 to 34	627	29.8%
	35 to 44	679	32.2%
	45 to 54	481	22.8%
	55 to 64	320	15.2%
	Total	2107	100.0%

그림 21. 범주가 제외된 테이블의 총계

모든 범주가 포함될 경우의 2,828에 비해 이 테이블의 총계는 2,107에 불과합니다. 테이블에 사용되는 범주만 총계에 포함됩니다. 모든 퍼센트는 데이터 파일의 총 케이스 수가 아니라 테이블에 사용된 총 케이스 수를 기준으로 하므로 이 경우에도 퍼센트 총계는 100%입니다.

총계의 출력 위치

기본적으로 총계는 총계화 중인 범주의 아래쪽에 표시됩니다. 총계화 중인 범주 위에 총계가 표시되도록 총계의 출력 위치를 변경할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
3. 총계 및 소계 표시 그룹에서 적용할 상위 범주를 선택합니다.
4. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Percent
Age category	Total	2107	100.0%
	25 to 34	627	29.8%
	35 to 44	679	32.2%
	45 to 54	481	22.8%
	55 to 64	320	15.2%

그림 22. 총계화된 범주 위에 표시된 총계

중첩 테이블의 총계

모든 중첩 수준에서 총계를 범주형 변수에 적용할 수 있으므로 다중 중첩 수준의 그룹 총계를 포함하는 테이블을 작성할 수 있습니다.

그룹 총계

다른 범주형 변수 내에 중첩된 범주형 변수의 총계를 그룹 총계라고 합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 성별을 캔버스 분할창의 연령범주 왼쪽으로 끌어다 놓습니다.
3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.

테이블을 작성하기 전에 총계를 총계화된 범주 아래로 다시 이동해 봅니다.

4. 총계 및 소계 표시 그룹에서 적용할 하위 범주를 선택합니다.
5. 적용을 클릭하여 설정을 저장한 다음 테이블 작성기로 돌아갑니다.
6. 확인을 클릭하여 테이블을 작성합니다.

			Count	Percent	
Gender	Male	Age category	25 to 34	276	29.3%
			35 to 44	309	32.8%
			45 to 54	221	23.5%
			55 to 64	136	14.4%
			Total	942	100.0%
	Female	Age category	25 to 34	351	30.1%
			35 to 44	370	31.8%
			45 to 54	260	22.3%
			55 to 64	184	15.8%
			Total	1165	100.0%

그림 23. 성별 범주 내의 연령범주 총계

테이블에 남자와 여자에 대한 두 그룹 총계가 표시됩니다.

총계

중첩 변수에 적용된 총계는 총계가 아니라 항상 그룹 총계입니다. 전체 테이블에 대한 총계가 필요하다면 가장 바깥쪽에 있는 중첩 수준의 변수에 총계를 적용할 수 있습니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창에서 성별을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
3. 범주 및 총계 대화 상자에서 총계 확인란을 선택합니다.
4. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

				Count	Percent
Gender	Male	Age category	25 to 34	276	29.3%
			35 to 44	309	32.8%
			45 to 54	221	23.5%
			55 to 64	136	14.4%
			Total	942	100.0%
	Female	Age category	25 to 34	351	30.1%
			35 to 44	370	31.8%
			45 to 54	260	22.3%
			55 to 64	184	15.8%
			Total	1165	100.0%
	Total	Age category	25 to 34	627	29.8%
			35 to 44	679	32.2%
			45 to 54	481	22.8%
			55 to 64	320	15.2%
Total			2107	100.0%	

그림 24. 중첩 테이블의 총계

총계는 2,828이 아니라 2,107에 불과합니다. 두 연령범주가 테이블에서 제외되어 해당 범주의 케이스가 총계에서 제외됩니다.

레이어 변수 총계

레이어 변수의 총계는 테이블에 개별 레이어로 표시됩니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 테이블 작성기에서 레이어를 클릭하여 레이어 목록을 표시합니다.
3. 캔버스 분할창의 행 영역에서 성별을 레이어 목록으로 끌어다 놓습니다.

참고: 성별에 대한 총계를 이미 지정했으므로 해당 총계를 지정할 필요가 없습니다. 차원 간 변수 이동은 해당 변수의 설정에 영향을 미치지 않습니다.

4. 확인을 클릭하여 테이블을 작성합니다.
5. 뷰어에서 테이블을 두 번 클릭하여 활성화합니다.
6. 레이어 드롭다운 목록에서 아래쪽 화살표를 클릭하여 레이어 전체 목록을 테이블에 표시합니다.

성별 남자, 성별 여자, 연령범주 등의 세 레이어가 있습니다.

레이어 총계의 출력 위치

레이어 변수 총계에서 총계의 출력 위치(위 또는 아래)에 따라 총계의 레이어 위치가 결정됩니다. 예를 들어, 레이어 변수 총계에 대해 적용할 상위 범주를 지정하면 총계 레이어가 첫 번째 레이어로 표시됩니다.

소개

변수의 범주 서브세트에 대한 소개를 포함할 수 있습니다. 예를 들어, 표본 설문조사에서 연령이 45세 미만과 45세 이상인 모든 반응자를 나타내는 연령범주에 대한 소개를 포함할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
5. 값 목록에서 **3.00**을 선택합니다.
6. 소개 정의 대화 상자를 표시하려면 소개 추가를 클릭합니다.
7. 레이블 텍스트 필드에서 소개 < 45를 입력합니다.
8. 그런 다음 계속을 클릭합니다.

처음 세 연령범주에 대한 소개가 포함된 행이 삽입됩니다.

9. 값 목록에서 **6.00**을 선택합니다.
10. 소개 정의 대화 상자를 표시하려면 소개 추가를 클릭합니다.
11. 레이블 텍스트 필드에서 소개 45+를 입력합니다.
12. 그런 다음 계속을 클릭합니다.

중요 참고: 소개를 정의하기 전에 총계 및 소개의 표시 위치(적용할 상위 범주 또는 적용할 하위 범주)를 선택해야 합니다. 표시 위치 변경은 현재 선택된 소개를 제외한 모든 소개에 영향을 미치고 소개에 포함된 범주도 함께 변경됩니다.

13. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	Subtotal < 45	1548
	45 to 54	481
	55 to 64	320
	65 or older	479
	Subtotal 45+	1280

그림 25. 연령범주의 소개

표시된 내용대로 소개 생성

총계와 마찬가지로 테이블에 포함된 범주에 따라 소개가 계산됩니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).

2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.

참고: 첫 번째 소계에 표시되는 값(레이블 아님)은 **1.00...3.00**인데 이는 목록에 있는 1과 3 사이의 모든 값이 소계에 포함됨을 나타냅니다.

3. 값 목록에서 **1.00**을 선택하거나 25 미만이라는 레이블을 클릭합니다.
4. 제외 목록 왼쪽의 화살표 키를 클릭합니다.

첫 번째 연령범주가 제외되고 첫 번째 소계에 대해 표시된 값이 **2.00...3.00**으로 변경됩니다. 이는 테이블에 포함된 범주를 기준으로 소계가 계산되므로 제외된 범주가 소계에 포함되지 않음을 나타냅니다. 예를 들어, 범주를 제외하면 해당 범주가 소계에서 자동으로 제외되므로 소계 계산의 기준이 되는 범주 없이 소계만 표시할 수 없습니다.

소계 범주 감추기

소계를 정의하는 범주를 표시하지 않고 소계만을 표시함으로써 기본 데이터에 영향을 주지 않고 범주를 효과적으로 “합칠” 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
5. 값 목록에서 **3.00**을 선택합니다.
6. 소계 정의 대화 상자를 표시하려면 소계 추가를 클릭합니다.
7. 레이블 텍스트 필드에서 45 미만을 입력합니다.
8. 테이블에서 소계 범주 감추기를 선택합니다.
9. 그런 다음 계속을 클릭합니다.

처음 세 연령범주에 대한 소계가 포함된 행이 삽입됩니다.

10. 값 목록에서 **6.00**을 선택합니다.
11. 소계 정의 대화 상자를 표시하려면 소계 추가를 클릭합니다.
12. 레이블 텍스트 필드에서 45 이상을 입력합니다.
13. 소계 범주 감추기를 선택합니다.
14. 그런 다음 계속을 클릭합니다.
15. 소계와 함께 총계를 포함하려면 표시 그룹에서 총계를 선택합니다.
16. 적용을 클릭합니다.

캔버스에서 소계는 표시되지만 소계를 정의하는 범주는 제외됩니다.

17. 확인을 클릭하여 테이블을 작성합니다.

		Count
Age category	Less than 45	1548
	45 or older	1280
Total		2828

그림 26. 소계와 총계만을 표시하는 테이블

레이어 변수 소계

총계와 마찬가지로 레이어 변수의 소계는 테이블에 개별 레이어로 표시됩니다. 따라서 소계는 범주로 취급됩니다. 각 범주는 테이블에서 개별 레이어이고 레이어 범주의 출력 순서는 소계 범주의 출력 위치를 비롯하여 범주 및 총계 대화 상자에 지정된 범주 순서에 따라 결정됩니다.

제 5 장 범주형 변수의 계산된 범주

통계표에 계산된 범주를 포함할 수 있습니다. 이 범주는 행, 열 또는 레이어 등 모든 차원의 중첩된 수준에서 동일한 변수의 범주에서 계산된 새로운 범주입니다. 예를 들어, 두 범주 간의 차이를 보여주는 계산된 범주를 포함할 수 있습니다.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

단순 계산된 범주

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
4. 값 목록에서 **3.00**을 선택합니다.
5. 계산된 범주 정의 대화 상자를 표시하려면 범주 추가를 클릭합니다.
6. 계산된 범주 레이블 텍스트 필드에서 45 미만을 입력합니다.
7. 범주 목록에서 **25 미만(1.00)**을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다. [1]이 표현식에 표시됩니다.
8. 대화 상자에서 더하기(+) 연산자 단추를 클릭하거나 키보드에서 + 키를 클릭합니다.
9. 범주 목록에서 **25 - 34(2.00)**을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
10. 대화 상자에서 더하기(+) 연산자 단추를 클릭하거나 키보드에서 + 키를 클릭합니다.
11. 범주 목록에서 **35 - 44(3.00)**을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
12. 그런 다음 계속을 클릭합니다.

처음 세 연령범주에 대한 소계가 포함된 행이 삽입됩니다.

13. 값 목록에서 **5.00**을 선택합니다.
14. 소계 정의 대화 상자를 표시하려면 소계 추가를 클릭합니다.
15. 레이블 텍스트 필드에서 65 미만을 입력합니다.
16. 그런 다음 계속을 클릭합니다.

처음 5개 범주에 대한 소계가 포함된 행이 삽입됩니다.

17. 테이블 작성기에서 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	Less than 45	1548
	45 to 54	481
	55 to 64	320
	Less than 65	2349
	65 or older	479

그림 27. 소계가 포함된 계산된 범주

이 테이블에는 계산된 범주(45 미만)와 소계(65 미만)이 포함됩니다. 소계에는 계산된 범주에 포함된 범주도 포함됩니다. 소계에서는 동일한 범주를 공유할 수 없으므로 소계만을 사용하여 동일한 테이블을 작성할 수 없습니다.

계산된 범주에서 범주 감추기

소계에서와 마찬가지로, 계산된 범주의 표현식에 사용되는 범주와 계산된 범주 자체가 표시되지 않도록 지정할 수 있습니다. 다음 예제는 이전 예제에서 작성됩니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.

3. 값 목록에서 45 미만 계산된 범주를 선택합니다.

4. 계산된 범주 정의 대화 상자를 표시하려면 편집을 클릭합니다.

5. 테이블에서 표현식에 사용되는 범주 감추기를 선택합니다.

6. 그런 다음 계속을 클릭합니다.

7. 값 목록에서 65 미만 소계를 선택합니다.

8. 소계 정의 대화 상자를 표시하려면 편집을 클릭합니다.

9. 테이블에서 소계 범주 감추기를 선택합니다.

10. 그런 다음 계속을 클릭합니다.

11. 테이블 작성기에서 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count
Age category	Less than 45	1548
	Less than 65	2349
	65 or older	479

그림 28. 소계 및 감춰진 범주가 있는 계산된 범주

이전 예제에서와 같이, 테이블에 계산된 범주 및 소계가 포함됩니다. 하지만 이 경우, 총계만 표시되도록 각 범주가 감춰져 있습니다.

계산된 범주의 소계 참조

계산된 범주의 표현식에 소계를 포함할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록의 고용상태를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 결혼상태를 열 영역으로 끌어다 놓습니다.
5. 캔버스 분할창에서 고용상태를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
6. 값 목록에서 2를 선택합니다.
7. 소계 정의 대화 상자를 표시하려면 소계 추가를 클릭합니다.
8. 레이블 텍스트 필드에서 근무를 입력합니다.
9. 테이블에서 소계 범주 감추기를 선택합니다.
10. 그런 다음 계속을 클릭합니다.

처음 2개 근무 상태 범주에 대한 소계가 포함된 행이 삽입됩니다.

11. 값 목록에서 8을 선택합니다.
12. 소계 정의 대화 상자를 표시하려면 소계 추가를 클릭합니다.
13. 레이블 텍스트 필드에서 휴직을 입력합니다.
14. 소계 범주 감추기를 선택합니다.
15. 그런 다음 계속을 클릭합니다.

기타 근무 상태 범주에 대한 소계가 포함된 행이 삽입됩니다.

16. 값 목록에서 휴직 소계를 선택합니다.
17. 계산된 범주 정의 대화 상자를 표시하려면 범주 추가를 클릭합니다.
18. 계산된 범주 레이블 텍스트 필드에서 근무/휴직을 입력합니다.
19. 총계 및 소계 목록에서 근무(근무 #1)를 선택하고 회살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.

20. 대화 상자에서 나누기(/) 연산자 단추를 클릭하거나 키보드에서 / 키를 클릭합니다.
21. 총계 및 소계 목록에서 휴직(휴직 #2)을 선택하고 회살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.

기본적으로 계산된 범주에서는 변수의 통계와 동일한 형식을 사용합니다(이 경우 자릿수). 계산된 범주의 표현식에서 나는 결과 값을 소수점 이하 자릿수로 표시하려 하지만 개수의 기본 형식에 소수점 이하 자릿수가 없기 때문에 형식을 변경해야 합니다.
22. 표시 형식 탭을 클릭합니다.
23. 소수점 자릿수 설정을 2로 변경합니다.
24. 그런 다음 계속을 클릭합니다.
25. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		Count	Count	Count	Count	Count
Labor force status	Working	916	64	330	67	494
	Not Working	429	219	116	26	169
	Working / Not Working	2.14	.29	2.84	2.58	2.92

그림 29. 소계 비율을 보여주는 계산된 범주

이 테이블에는 2개의 소계와 1개의 계산된 범주가 포함됩니다. 계산된 범주에 소계 비율이 표시되므로 각 소계에 나타내는 그룹을 쉽게 비교할 수 있습니다. 다른 그룹과 비교할 때 근무하는 사별 반응자 대 휴직한 사별 반응자의 비율이 훨씬 낮습니다. 또한 기혼 반응자의 비율이 약간 낮습니다. 가정에서 아이를 돌보기 위해 배우자가 직장을 떠났기 때문일 수 있습니다.

계산된 범주를 사용하여 비전체적인 소계 표시

소계는 전체적입니다. 즉, 테이블의 모든 소계에는 테이블에서 해당 값 위치 위 또는 아래의 모든 값이 포함됩니다. 반면에, 계산된 범주는 전체적이 아니며 테이블의 범주를 혼합하여 합계를 낼 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 테이블 작성기의 변수 목록에서 자신을 진보적 또는 보수적으로 생각을 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 캔버스 분할창에서 자신을 진보적 또는 보수적으로 생각을 마우스 오른쪽 단추로 클릭하고 팝업 컨텍스트 메뉴에서 범주 및 총계를 선택합니다.
5. 값 목록에서 3을 선택합니다.
6. 계산된 범주 정의 대화 상자를 표시하려면 범주 추가를 클릭합니다.

7. 계산된 범주 레이블 텍스트 필드에서 진보 소계를 입력합니다. 텍스트 앞에 4개의 공백이 있습니다. 이러한 공백은 결과 테이블에서 들여쓰기에 사용됩니다.
8. 범주 목록에서 매우 진보적(1)을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
9. 대화 상자에서 더하기(+) 연산자 단추를 클릭하거나 키보드에서 + 키를 클릭합니다.
10. 범주 목록에서 진보적(2)을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
11. 대화 상자에서 더하기(+) 연산자 단추를 클릭하거나 키보드에서 + 키를 클릭합니다.
12. 범주 목록에서 약간 진보적(3)을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
13. 계속을 클릭합니다.

진보 범주에 대한 소계가 포함된 행이 삽입됩니다.

14. 값 목록에서 7을 선택합니다.
15. 계산된 범주 정의 대화 상자를 표시하려면 범주 추가를 클릭합니다.
16. 계산된 범주 레이블 텍스트 필드에서 보수 소계를 입력합니다. 텍스트 앞에 4개의 공백이 있습니다. 이러한 공백은 결과 테이블에서 들여쓰기에 사용됩니다.
17. 범주 목록에서 약간 보수적(5)을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
18. 대화 상자에서 더하기(+) 연산자 단추를 클릭하거나 키보드에서 + 키를 클릭합니다.
19. 범주 목록에서 보수적(6)을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
20. 대화 상자에서 더하기(+) 연산자 단추를 클릭하거나 키보드에서 + 키를 클릭합니다.
21. 범주 목록에서 매우 보수적(7)을 선택하고 화살표 단추를 클릭하여 계산된 범주의 표현식 입력란에 복사합니다.
22. 계속을 클릭합니다.

보수 범주에 대한 소계가 포함된 행이 삽입됩니다.

23. 테이블 작성기에서 적용 을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count
Think of self as liberal or conservative	Extremely liberal	64
	Liberal	357
	Slightly liberal	351
	Liberal Subtotal	772
	Moderate	986
	Slightly conservative	432
	Conservative	415
	Extremely conservative	86
	Conservative Subtotal	933

그림 30. 비전체적인 소계를 표시하는 계산된 범주

이 테이블에는 테이블에 표시된 모든 범주를 포함하지 않는 2개의 계산된 범주가 포함됩니다. 보통 범주는 두 계산된 범주에 포함되지 않습니다. 소계는 전체적이므로 소계를 사용하여 동일한 테이블을 작성할 수 없습니다.

제 6 장 공유 범주가 있는 변수 테이블

설문조사에 포함된 많은 질문은 공통적인 반응 세트를 가지기도 합니다. 예를 들어, 표본 설문조사에는 공공 기관 및 사설 기관과 해당 서비스의 신뢰도에 대한 여러 변수와 1 = 높음, 2 = 보통, 3 = 낮음과 같은 동일한 반응 범주가 포함됩니다. 누적을 사용하면 관련 있는 변수를 같은 테이블에 표시한 다음 공유 반응 범주를 테이블의 열에 표시할 수 있습니다. 또한 계산된 범주를 사용하는 경우, 계산된 범주의 레이블 및 표현식이 모든 변수에서 동일하다는 조건 하에 이러한 기능을 사용할 수 있습니다.

	A great deal	Only some	Hardly any
Confidence in banks & financial institutions	490	1068	306
Confidence in education	511	1055	315
Confidence in major companies	500	1078	243
Confidence in medicine	844	864	167
Confidence in press	176	878	808
Confidence in television	196	936	744

그림 31. 공유 범주가 있는 변수 테이블

참고: 이전 버전의 테이블 사용자 정의에서는 이것을 "빈도 통계표"라고 했습니다.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 설정합니다.

개수표

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기의 변수 목록에서 은행 및 금융기관에 대한 신뢰도를 클릭한 다음 Shift 키를 누른 채 TV에 대한 신뢰도를 클릭하여 "신뢰도" 변수를 모두 선택합니다. (참고: 여기에서는 변수 레이블이 파일순이 아니라 문자순으로 변수 목록에 표시된다고 가정합니다.)
3. 여섯 개의 신뢰도 변수를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

변수가 행 차원에 누적됩니다. 기본적으로 각 변수의 범주 레이블도 행에 표시되기 때문에 테이블이 아주 길고 좁아집니다(6개의 변수 x 3개의 범주 = 18개 행). 그러나 여섯 개 변수의 범주 레이블(값 레이블)이 동일하므로 범주 레이블을 열 차원에 놓을 수 있습니다.

4. 범주 위치 드롭다운 목록에서 **행 레이블을 열로 표시**를 선택합니다.

이제 테이블에는 누적변수마다 하나씩 총 여섯 개의 행만 있고 각 변수에 동일하게 정의된 범주가 테이블의 열이 됩니다.

5. **개수**라는 요약 통계 레이블은 실제로 필요하지 않으므로 테이블을 만들기 전에 요약 통계의 위치에서 숨김 확인란을 선택합니다.
6. **확인**을 클릭하여 테이블을 작성합니다.

	A great deal	Only some	Hardly any
Confidence in banks & financial institutions	490	1068	306
Confidence in education	511	1055	315
Confidence in major companies	500	1078	243
Confidence in medicine	844	864	167
Confidence in press	176	878	808
Confidence in television	196	936	744

그림 32. 공유 범주 레이블을 열로 표시한 누적 행 변수에 대한 테이블

행에 변수를 표시하고 열에 범주를 표시하는 대신 열에 변수를 누적하고 행에 범주를 표시한 테이블을 만들 수 있습니다. 변수보다 범주가 많을 경우에 이 테이블이 적합하지만 이 예제에서는 범주보다 변수가 더 많습니다.

퍼센트 테이블

변수를 행에 누적하고 범주를 열에 표시한 테이블의 경우 가장 적합하거나 이해하기 쉬운 표시 퍼센트는 행 퍼센트입니다 (변수를 열에 누적하고 범주를 행에 표시한 테이블의 경우에는 열 퍼센트).

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 신뢰도 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 **요약 통계**를 선택합니다.
3. 통계 목록에서 **행 N %**를 선택하고 회살표 단추를 클릭하여 출력 목록으로 옮깁니다.
4. 출력 목록에서 **개수** 행의 셀을 선택하고 회살표 단추를 클릭하여 통계 목록으로 다시 옮겨 출력 목록에서 제거합니다.
5. **모든 항목에 적용**을 클릭하여 요약 통계 변경 사항을 테이블의 누적변수 모두에 적용합니다.

참고: 테이블 미리보기가 이 그림과 틀리다면 아마도 모든 항목에 적용이 아닌 선택한 항목에 적용을 클릭하여 선택한 변수에만 새 요약 통계가 적용되었기 때문일 수 있습니다. 이러한 경우 범주별로 두 개의 열

이 표시됩니다. 한 열에는 선택한 변수에 대한 행 퍼센트 자리표시자가 표시되고 다른 열에는 기타 모든 변수에 대한 개수 자리표시자가 표시됩니다. 정확히 이와 같은 테이블이 생성되겠지만 이 테이블은 이 예제에서 필요한 테이블이 아닙니다.

6. 확인을 클릭하여 테이블을 작성합니다.

	A great deal	Only some	Hardly any
Confidence in banks & financial institutions	26.3%	57.3%	16.4%
Confidence in education	27.2%	56.1%	16.7%
Confidence in major companies	27.5%	59.2%	13.3%
Confidence in medicine	45.0%	46.1%	8.9%
Confidence in press	9.5%	47.2%	43.4%
Confidence in television	10.4%	49.9%	39.7%

그림 33. 열에 범주가 표시되어 있고 행에 행 퍼센트와 함께 누적된 변수 테이블

참고: 공유 범주가 있는 변수 테이블에 요약 통계를 제한 없이 포함시킬 수 있지만 이 예제에서는 요약 통계의 적용을 간단하게 알 수 있도록 한 번에 하나만 표시합니다.

총계 및 범주 제어

테이블의 모든 변수가 동일한 순서로 표시된 동일한 범주를 가지지만 변수와 다른 차원에 범주를 표시한 테이블을 만들 수 있습니다. 이를 위해서는 총계, 소계 및 기타 범주를 조정해야 합니다. 따라서 범주 및 총계 대화 상자에서 변경한 사항은 테이블에서 범주를 공유하는 모든 변수에 적용되어야 합니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 첫 번째 신뢰도 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
3. 범주 및 총계 대화 상자에서 총계를 선택한 다음 적용을 클릭합니다.

이 그림에서 우선 범주 레이블이 열에서 행으로 다시 옮겨졌음을 알 수 있습니다. 또한 범주 위치 제어가 비활성되어 있음을 알 수 있습니다. 이는 변수가 더 이상 정확히 같은 "범주 세트"를 공유하지 않기 때문입니다. 변수 중 하나에 총계 범주가 포함되어 있습니다.

4. 캔버스 분할창에서 신뢰도 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 모든 행 변수 선택을 선택합니다. 또는 Ctrl 키를 누른 채 각각의 누적변수를 클릭하여 변수를 모두 선택합니다. 이를 위해 창을 아래로 스크롤하거나 테이블 작성기 창을 확장해야 할 수도 있습니다.
5. 정의에서 범주 및 총계를 클릭합니다.
6. 범주 및 총계 대화 상자에서 총계가 아직 선택되어 있지 않은 경우 지금 선택한 다음 적용을 클릭합니다.
7. 이제 모든 변수에 총계 범주가 추가되었으므로 범주 위치 드롭다운 목록을 다시 사용할 수 있습니다. 따라서 행 레이블을 열로 표시를 선택할 수 있습니다.
8. 확인을 클릭하여 테이블을 작성합니다.

	A great deal	Only some	Hardly any	Total
Confidence in banks & financial institutions	26.3%	57.3%	16.4%	100.0%
Confidence in education	27.2%	56.1%	16.7%	100.0%
Confidence in major companies	27.5%	59.2%	13.3%	100.0%
Confidence in medicine	45.0%	46.1%	8.9%	100.0%
Confidence in press	9.5%	47.2%	43.4%	100.0%
Confidence in television	10.4%	49.9%	39.7%	100.0%

그림 34. 열에 범주와 총계가 표시되어 있고 행에 행 퍼센트와 함께 변수가 누적되어 있는 테이블

공유 범주가 있는 테이블에 중첩

범주 레이블을 반대 차원에 표시하려면 중첩된 테이블에서 공유 범주가 있는 누적변수가 해당 차원의 가장 안쪽 중첩 수준에 있어야 합니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록에서 성별을 행 영역의 왼쪽으로 끌어다 놓습니다.

범주를 공유하는 누적변수는 이제 테이블 미리보기에서 성별 범주 내에 중첩됩니다.

3. 이제 성별을 테이블 미리보기에서 누적된 신뢰도 변수의 오른쪽으로 끌어다 놓습니다.

마찬가지로 범주 레이블이 행 차원으로 다시 옮겨졌고 범주 위치 제어도 비활성되어 있습니다. 이제 하나의 누적변수에 성별이 중첩되어 있지만 다른 누적변수에는 중첩 변수가 없습니다. 성별을 각각의 누적변수에 중첩 변수로 추가할 수 있지만 행 레이블을 열로 옮기면 공유 범주가 있는 누적변수의 범주 레이블이 아닌 성별의 범주 레이블이 열에 표시됩니다. 이는 이제 성별이 가장 안쪽에 있는 중첩 변수이고 범주 위치 변경은 항상 가장 안쪽 중첩 변수에 적용되기 때문입니다.

제 7 장 요약 통계

요약 통계는 범주형 변수의 단순 개수부터 척도변수의 평균 표준 오차와 같은 산포 측도까지 모든 것을 포함합니다. 테이블 사용자 정의 대화 상자의 검정 통계량 탭에서 사용할 수 있는 유의성 검증은 포함되지 않습니다. 자세한 정보는 71 페이지의 『검정 통계량』 주제를 참조하십시오.

범주형 변수 및 다중 응답 세트에 대한 요약 통계에는 다음과 같은 개수 및 다양한 퍼센트 계산이 포함됩니다.

- 행 퍼센트
- 열 퍼센트
- 하위 표 퍼센트
- 표 퍼센트
- 유효수 퍼센트

다음은 범주형 변수에 사용할 수 있는 요약 통계뿐만 아니라 척도변수에 대한 요약 통계와 범주형 변수에 대한 사용자 정의 총계 요약에도 포함됩니다.

- 평균
- 중위수
- 백분위수
- 합계
- 표준 편차
- 범위
- 최소값 및 최대값

추가 요약 통계를 다중 응답 세트에 사용할 수 있습니다. 요약 통계의 전체 목록도 사용할 수 있습니다. 자세한 정보는 6 페이지의 『요약 통계』 주제를 참조하십시오.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 설정합니다.

요약 통계 소스 변수

사용할 수 있는 요약 통계는 요약 통계 소스 변수의 측정 수준에 따라 다릅니다. 요약 통계의 소스(요약 통계 기준 변수)는 다음에 의해 결정됩니다.

- 측정 수준. 테이블 또는 누적표의 테이블 섹션에 척도변수가 있는 경우 요약 통계는 척도변수를 기준으로 합니다.
- 변수 선택 순서. 범주형 변수에 대한 기본값 통계 소스 차원(행 또는 열)은 캔버스 분할창으로 변수를 끌어서 놓는 순서를 기준으로 합니다. 예를 들어, 변수를 행 영역으로 먼저 끌면 행 차원이 기본값 통계 소스 차원이 됩니다.
- 중첩. 범주형 변수의 경우 요약 통계는 통계 소스 차원에서 가장 안쪽에 있는 변수를 기준으로 합니다.

누적표에는 다중 요약 통계 소스 변수(척도변수와 범주형 변수 모두)가 있을 수 있지만 각 테이블 섹션에는 요약 통계 소스가 하나만 있습니다.

범주형 변수의 요약 통계 소스

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록의 연령범주를 캔버스 분할창의 행 영역으로 끌어서 놓습니다.
3. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다. 연령범주는 이 테이블의 유일한 변수이기 때문에 통계 소스 변수가 됩니다.
4. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
5. 선택한 항목에 적용을 클릭합니다.
6. 테이블 작성기에서 인터넷으로부터 뉴스언음을 캔버스 분할창의 연령범주 오른쪽으로 끌어서 놓습니다.
7. 캔버스 분할창에서 연령범주를 마우스 오른쪽 단추로 다시 클릭합니다. 이제 연령범주가 통계 소스 차원의 가장 안쪽에 있는 중첩 변수가 아니기 때문에 팝업 메뉴에서 요약 통계 항목이 비활성화됩니다.
8. 캔버스 분할창에서 인터넷으로부터 뉴스언음을 마우스 오른쪽 단추로 클릭합니다. 이제 이 변수가 통계 소스 차원의 가장 안쪽에 있는 중첩 변수이기 때문에 요약 통계 소스 변수가 되고 요약 통계 항목이 활성화됩니다. 테이블에 한 차원행만 있기 때문에 행이 통계량 소스 차원이 됩니다.
9. 캔버스 분할창의 행 영역에서 인터넷으로부터 뉴스언음을 열 영역으로 끌어서 놓습니다.
10. 캔버스 분할창에서 인터넷으로부터 뉴스언음을 마우스 오른쪽 단추로 다시 클릭합니다. 이제 이 변수가 더 이상 통계 소스 차원에 없기 때문에 팝업 메뉴에서 요약 통계 항목이 비활성화됩니다.

테이블을 작성할 때 먼저 변수가 놓인 차원이 범주형 변수의 기본값 통계 소스 차원이 되기 때문에 연령범주가 다시 통계 소스 변수가 됩니다. 이 예제에서는 가장 먼저 행 차원에 변수를 놓았습니다. 따라서 행 차원이 기본값 통계 소스 차원이 되고, 연령범주는 해당 차원의 유일한 변수이기 때문에 통계 소스 변수가 됩니다.

척도변수의 요약 통계 소스

1. 하루 TV시청시간 척도변수를 캔버스 분할창의 행 영역에 있는 연령범주 왼쪽으로 끌어서 놓습니다.

여기서 개수 및 열 $N\%$ 요약이 평균으로 대체되었음을 알 수 있습니다. 또한 캔버스 분할창에서 하루 TV 시청시간을 마우스 오른쪽 단추로 클릭하면 이 변수가 요약 통계 소스 변수가 됩니다. 척도변수가 있는 테이블에서 척도변수는 중첩 수준이나 차원에 관계 없이 항상 통계 소스 변수이고, 척도변수의 기본값 요약 통계는 평균입니다.

2. 하루 TV시청시간을 행 영역에서 인터넷으로 부터 뉴스얻음 위의 열 영역으로 끌어다 놓습니다.
3. 하루 TV시청시간을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다. 이 변수는 다른 차원으로 이동하더라도 통계 소스 변수입니다.
4. 요약 통계 대화 상자의 출력 목록에서 평균에 대한 형식 셀을 클릭하고 형식 드롭다운 목록에서 nnnn을 선택합니다. 이 선택 항목을 찾으려면 목록을 위로 스크롤해야 할 수 있습니다.
5. 소수점이하자리 셀에 2를 입력합니다.
6. 선택한 항목에 적용을 클릭합니다.

캔버스 분할창의 테이블 미리보기에 평균값이 소수점 2자리로 표시됩니다.

7. 확인을 클릭하여 테이블을 작성합니다.

		Hours per day watching TV	
		Get news from internet	
		No	Yes
		Mean	Mean
Age category	Less than 25	3.54	2.12
	25 to 34	3.42	2.14
	35 to 44	3.00	2.01
	45 to 54	2.83	2.06
	55 to 64	3.24	2.37
	65 or older	3.82	2.33

그림 35. 교차 분석표 범주형 변수 내에 요약된 척도변수

누적변수

누적표에 다중 통계 소스 변수가 포함될 수 있고 각각의 통계 소스 변수에 대해 서로 다른 요약 통계를 지정할 수 있기 때문에 누적표에 요약 통계를 지정할 경우 몇 가지 고려할 사항이 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 변수 목록에서 인터넷으로부터 뉴스얻음을 클릭한 다음 변수 목록에서 TV로부터 뉴스얻음을 Shift를 누른 채 마우스로 클릭하여 모든 "뉴스" 변수를 선택합니다. (참고: 여기에서는 변수 레이블이 파일순이 아니라 문자순으로 변수 목록에 표시된다고 가정합니다.)
4. 다섯 개의 뉴스 변수를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

다섯 개의 뉴스 변수가 행 차원에 누적됩니다.

5. 캔버스 분할창에서 인터넷으로부터 뉴스얻음을 클릭하여 해당 변수만 선택합니다.

6. 인터넷으로부터 뉴스언음을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
7. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다. 화살표를 사용하여 선택한 통계를 통계 목록에서 출력 목록으로 이동하거나, 선택한 통계를 통계 목록에서 출력 목록으로 끌어다 놓을 수 있습니다.
8. 그런 다음 선택한 항목에 적용을 클릭합니다.

열 퍼센트에 대해 열이 추가되지만, 캔버스 분할창의 테이블 미리보기에는 한 변수에 대해서만 열 퍼센트가 표시됩니다. 누적표에 다중 통계 소스 변수가 있고 각 통계 소스 변수가 서로 다른 요약 통계를 가질 수 있기 때문입니다. 그러나, 이 예제에서는 모든 변수에 대해 동일한 요약 통계를 표시할 것입니다.

9. 캔버스 분할창에서 신문으로부터 뉴스언음을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
10. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
11. 그런 다음 모든 항목에 적용을 클릭합니다.

테이블 미리보기에 모든 누적변수에 대해 열 퍼센트가 표시됩니다.

범주형 변수의 사용자 정의 총계 요약 통계

범주형 통계 소스 변수의 경우 변수 범주에 표시되는 통계와 다른 사용자 정의 총계 요약 통계를 포함할 수 있습니다. 예를 들어, 순서변수의 경우 각 범주에 대한 퍼센트와 사용자 정의 총계 요약 통계에 대한 평균 또는 중위수를 표시할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 변수 목록에서 출판물에 대한 신뢰도를 클릭한 다음 TV에 대한 신뢰도를 Ctrl 클릭하여 두 변수를 모두 선택합니다.
4. 두 변수를 캔버스 분할창의 행 영역으로 끌어다 놓습니다. 행 차원에 두 변수가 누적됩니다.
5. 캔버스 분할창에서 두 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 모든 행 변수 선택을 선택합니다. 두 변수가 이미 모두 선택되어 있을 수도 있습니다.
6. 변수를 마우스 오른쪽 단추로 다시 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
7. 범주 및 총계 대화 상자에서 총계 확인란을 선택한 다음 적용을 클릭합니다.

캔버스 분할창의 테이블 미리보기에 두 변수에 대한 총계 행이 표시됩니다. 사용자 정의 총계 요약 통계를 표시하려면 테이블에 총계 및/또는 소계를 지정해야 합니다.

8. 캔버스 분할창에서 두 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
9. 요약 통계 대화 상자에서 출력 목록의 개수를 클릭하고 화살표를 클릭하여 통계 목록으로 이동합니다. 그러면 개수가 출력 목록에서 제거됩니다.
10. 통계 목록에서 열 N %를 클릭하고 화살표 키를 눌러 출력 목록으로 이동합니다.

11. 총계 및 소계에 대한 사용자 정의 요약 통계 확인란을 선택합니다.
12. 사용자 정의 요약 출력 목록에서 개수를 클릭하고 화살표를 클릭하여 사용자 정의 요약 통계 목록으로 이동합니다. 그러면 개수가 출력 목록에서 제거됩니다.
13. 사용자 정의 요약 통계 목록에서 평균을 클릭하고 화살표를 클릭하여 사용자 정의 요약 출력 목록으로 이동합니다.
14. 출력 목록에서 평균에 대한 형식 셀을 클릭하고 형식 드롭다운 목록에서 **nnnn**을 선택합니다. 이 선택 항목을 찾으려면 목록을 위로 스크롤해야 할 수 있습니다.
15. 소수점이하자리 셀에 2를 입력합니다.
16. 모든 항목에 적용을 클릭하여 테이블의 두 변수 모두에 해당 설정을 적용합니다.

사용자 정의 총계 요약 통계에 새 열이 추가되었습니다. 캔버스 분할창의 미리보기에 빈 셀이 많이 있는 테이블이 표시되기 때문에 새 열이 필요하지 않을 수도 있습니다.

17. 테이블 작성기에서 요약 통계의 위치 드롭다운 목록에서 행을 선택합니다.

모든 요약 통계가 행 차원으로 이동되고 모든 요약 통계가 테이블에 한 열로 표시됩니다.

18. 확인을 클릭하여 테이블을 작성합니다.

Confidence in press	A great deal	Column N %	9.5%
	Only some	Column N %	47.2%
	Hardly any	Column N %	43.4%
	Total	Mean	2.34
Confidence in television	A great deal	Column N %	10.4%
	Only some	Column N %	49.9%
	Hardly any	Column N %	39.7%
	Total	Mean	2.29

그림 36. 사용자 정의 총계 요약 통계가 있는 범주형 변수

범주 값 표시

이전 테이블에는 기본 범주 값을 모를 경우 평균값을 해석하기가 어렵다는 문제가 하나 있습니다. 평균 2.34는 높음과 보통 사이에 있습니까? 보통과 낮음 사이에 있습니까?

테이블 사용자 정의에서는 이 문제를 직접 해결할 수 없지만 보다 일반적인 방법으로 해결할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

편집 > 옵션...

2. 옵션 대화 상자에서 결과 레이블 탭을 클릭합니다.
3. 피벗 테이블 레이블 붙이기의 변수값의 레이블 표시는 다음과 같이 나타남 드롭다운 목록에서 값과 레이블을 선택합니다.
4. 확인을 클릭하여 이 설정을 저장합니다.

5. 테이블 작성기를 열고(분석 메뉴, 표, 테이블 사용자 정의) 확인을 클릭하여 테이블을 다시 작성합니다.

Confidence in press	1 A great deal	Column N %	9.5%
	2 Only some	Column N %	47.2%
	3 Hardly any	Column N %	43.4%
	Total	Mean	2.34
Confidence in television	1 A great deal	Column N %	10.4%
	2 Only some	Column N %	49.9%
	3 Hardly any	Column N %	39.7%
	Total	Mean	2.29

그림 37. 변수 범주에 표시되는 값과 레이블

범주 값으로 인해 평균 2.34가 보통과 낮음 사이에 있다는 것이 확실해집니다. 테이블에 범주 값을 표시하면 평균과 같은 사용자 정의 총계 요약 통계 값을 훨씬 쉽게 해석할 수 있습니다.

이 표시 설정은 모든 프로시저에서 모든 피벗 테이블 결과에 적용되고 변경하지 않는 한 세션 전체에 지속되는 광역 설정입니다. 값 레이블만 표시하도록 설정을 다시 변경하려면 다음과 같이 하십시오.

6. 메뉴에서 다음을 선택합니다.

편집 > 옵션...

7. 옵션 대화 상자에서 결과 레이블 탭을 클릭합니다.

8. 피벗 테이블 레이블 붙이기의 변수값의 레이블 표시는 다음과 같이 나타남 드롭다운 목록에서 레이블을 선택합니다.

9. 확인을 클릭하여 이 설정을 저장합니다.

제 8 장 척도변수 요약

척도변수 요약

척도변수에 대해 다양한 요약 통계를 사용할 수 있습니다. 척도변수에서는 범주형 변수에서 사용할 수 있는 개수 및 퍼센트뿐만 아니라 다음과 같은 요약 통계도 사용할 수 있습니다.

- 평균
- 중위수
- 백분위수
- 합계
- 표준 편차
- 범위
- 최소값 및 최대값

자세한 정보는 9 페이지의 『척도변수 및 범주형 사용자 정의 총계의 요약 통계』 주제를 참조하십시오.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 지정합니다.

누적 척도변수

척도변수를 테이블에 누적하면 동일한 테이블에서 여러 척도변수를 요약할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기의 변수 목록에서 반응자 연령을 클릭하고 Ctrl 키를 누른 채 피교육년수, 하루 TV 시청 시간을 차례로 클릭하여 세 개의 변수를 모두 선택합니다.
3. 선택한 세 개의 변수를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

세 개의 변수가 행 차원에 누적됩니다. 세 개의 변수가 모두 척도변수이므로 범주는 표시되지 않으며 기본 값 요약 통계는 평균입니다.

4. 확인을 클릭하여 테이블을 작성합니다.

	Mean
Age of respondent	46
Highest year of school completed	13
Hours per day watching TV	3

그림 38. 누적 척도변수의 평균값 테이블

다중 요약 통계

척도변수의 경우 기본적으로 평균이 표시되지만 척도변수에 대해 다른 요약 통계를 선택할 수 있고 둘 이상의 요약 통계를 표시할 수도 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 세 개의 척도변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
3. 요약 통계 대화 상자에서 통계 목록의 중위수를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다. 화살표를 사용하여 선택한 통계를 통계 목록에서 출력 목록으로 이동할 수도 있고 선택한 통계를 통계 목록에서 출력 목록으로 끌어다 놓을 수도 있습니다.
4. 출력 목록에서 중위수의 형식 셀을 클릭하고 형식 드롭다운 목록에서 **nnnn**을 선택합니다.
5. 소수점이하자리 셀에 1을 입력합니다.
6. 출력 목록의 평균에 대해서도 동일하게 변경합니다.
7. 모든 항목에 적용을 클릭하여 세 개의 척도변수 모두에 변경 사항을 적용합니다.
8. 테이블 작성기에서 확인을 클릭하여 테이블을 작성합니다.

	Mean	Median
Age of respondent	45.6	42.0
Highest year of school completed	13.3	13.0
Hours per day watching TV	2.9	2.0

그림 39. 누적 척도변수 테이블에 표시된 평균 및 중위수

개수, 유효수 및 결측

평균과 같은 요약 통계 계산에 사용된 케이스 수를 표시하는 것이 유용할 때가 많으며 이러한 정보를 개수라는 요약 통계에서 제공한다고 생각해도 틀린 것은 아닙니다. 하지만 결측값이 있는 경우 빈도에서 정확한 케이스 기준을 제공하지 못합니다. 정확한 케이스 기준을 구하려면 유효수를 사용합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).

2. 캔버스 분할창의 테이블 미리보기에서 세 개의 척도변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
3. 요약 통계 대화 상자에서 통계 목록의 개수를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
4. 그런 다음 통계 목록에서 유효수를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
5. 모든 항목에 적용을 클릭하여 세 개의 척도변수 모두에 변경 사항을 적용합니다.
6. 테이블 작성기에서 확인을 클릭하여 테이블을 작성합니다.

	Mean	Median	Count	Valid N
Age of respondent	45.6	42.0	2832	2828
Highest year of school completed	13.3	13.0	2832	2820
Hours per day watching TV	2.9	2.0	2832	2337

그림 40. 개수와 유효수 비교

세 개의 척도변수의 개수는 2,832로 모두 동일합니다. 이 값은 데이터 파일의 총 케이스 수와 일치합니다. 모든 범주형 변수 내에 척도변수가 서로 중첩되지 않으므로 이 개수 값과 데이터 파일의 총 케이스 수가 서로 일치하는 것입니다.

이와 달리 척도변수의 유효수는 서로 다르며 특히, 하루 TV 시청 시간의 경우에는 개수와 유효수가 매우 다릅니다. 이는 이 변수에 많은 결측값이 있기 때문입니다. 이 변수에 기록된 값이 없거나 결측 데이터를 표시하기로 정의된 값(예: 남자가 임신한 경우 해당없음을 표시하는 코드 99)이 있는 케이스입니다.

7. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
8. 캔버스 분할창의 테이블 미리보기에서 세 개의 척도변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
9. 요약 통계 대화 상자에서 출력 목록의 유효수를 선택하고 화살표 키를 눌러 통계 목록으로 다시 옮겨 출력 목록에서 제거합니다.
10. 출력 목록에서 개수를 선택하고 화살표를 클릭하여 통계 목록으로 다시 옮겨 출력 목록에서 제거합니다.
11. 통계 목록에서 결측을 선택하고 화살표 키를 눌러 출력 목록에 추가합니다.
12. 모든 항목에 적용을 클릭하여 세 개의 척도변수 모두에 변경 사항을 적용합니다.
13. 테이블 작성기에서 확인을 클릭하여 테이블을 작성합니다.

	Mean	Median	Missing
Age of respondent	45.6	42.0	4
Highest year of school completed	13.3	13.0	12
Hours per day watching TV	2.9	2.0	495

그림 41. 척도 요약 통계 테이블에 표시된 결측 수

이제 테이블에 각 척도변수의 결측값 수가 표시됩니다. 이 테이블을 보면 다른 두 변수에 비해 하루 TV 시청 시간의 결측값 수가 상대적으로 매우 많음을 확인할 수 있습니다. 변수 요약값에 따라 최종 통계 분석을 내리기 전에 이 기능을 통해 해당 결측을 파악해 볼 수 있습니다.

변수마다 다른 요약

다중 요약 통계를 표시할 수 있을 뿐만 아니라 누적표에 척도변수마다 다른 요약 통계를 표시할 수 있습니다. 예를 들어, 앞의 테이블에서 세 개의 변수 중 한 변수의 결측이 매우 많음을 확인했으므로 해당 변수의 결측값 수만 표시해 볼 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 반응자연령을 클릭하고 Ctrl 키를 누른 채 피교육년수를 클릭하여 두 변수를 선택합니다.
3. 선택한 두 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
4. 요약 통계 대화 상자에서 출력 목록의 결측을 선택하고 화살표 키를 눌러 통계 목록으로 다시 옮겨 출력 목록에서 제거합니다.
5. 선택한 항목에 적용을 클릭하여 선택한 두 변수에만 변경 사항을 적용합니다.

테이블에서 다른 두 변수의 데이터 셀에 자리표시자가 없기 때문에 하루 TV시청시간의 결측값 수만 표시됨을 알 수 있습니다.

6. 확인을 클릭하여 테이블을 작성합니다.

	Mean	Median	Missing
Age of respondent	45.6	42.0	
Highest year of school completed	13.3	13.0	
Hours per day watching TV	2.9	2.0	495

그림 42. 변수마다 다른 요약 통계가 표시된 테이블

이 테이블은 원하는 정보는 제공하지만 레이아웃 때문에 테이블을 잘못 해석할 수 있습니다. 즉, 결측 열에서 셀이 비어있는 변수에 결측이 없다고 해석할 수도 있습니다.

7. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
8. 테이블 작성기에서 요약 통계의 위치 드롭다운 목록에서 행을 선택합니다.
9. 확인을 클릭하여 테이블을 작성합니다.

Age of respondent	Mean	45.6
	Median	42.0
Highest year of school completed	Mean	13.3
	Median	13.0
Hours per day watching TV	Mean	2.9
	Median	2.0
	Missing	495

그림 43. 행 차원에 모두 표시된 요약 통계 및 변수

이제 테이블에 한 변수의 결측값 수만 표시된다는 사실을 분명히 알 수 있습니다.

범주별 그룹 요약

범주형 변수를 그룹화 변수로 사용하여 범주형 변수의 범주별로 정의된 그룹에 척도변수 요약 표시할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록에서 성별을 캔버스 분할창의 열 영역으로 끌어다 놓습니다.

캔버스 분할창의 테이블 미리보기에서 성별을 마우스 오른쪽 단추로 클릭하면 팝업 메뉴에서 요약 통계를 사용할 수 없음을 알 수 있습니다. 척도변수가 있는 테이블의 경우 척도변수가 항상 통계 소스 변수가 되기 때문입니다.

3. 확인을 클릭하여 테이블을 작성합니다.

		Gender	
		Male	Female
Age of respondent	Mean	44.6	46.3
	Median	42.0	43.0
Highest year of school completed	Mean	13.4	13.2
	Median	13.0	13.0
Hours per day watching TV	Mean	2.8	2.9
	Median	2.0	2.0
	Missing	213	282

그림 44. 범주형 열 변수를 사용한 그룹 척도 요약

이 테이블을 보면 남자와 여자에 대한 평균값(평균 및 중위수)을 쉽게 비교할 수 있으며 이들 사이의 차이가 그리 크지 않음을 분명히 알 수 있습니다. 이러한 정보는 유용할 수는 있지만 매우 흥미로운 것은 아닙니다.

다중 그룹화 변수

행 및 열 범주형 그룹화 변수를 사용하거나 중첩을 통해 그룹을 더욱 세분할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록에서 인터넷으로부터 뉴스얻음을 캔버스 분할창의 가장 왼쪽 행 영역으로 끌어다 놓습니다. 세 개의 척도변수 모두 이 변수에 중첩되도록 위치를 지정합니다.

위의 두 번째 예제와 같은 표모양도 필요한 경우가 있을 수 있지만 여기서 원하는 표모양은 아닙니다.

3. 확인을 클릭하여 테이블을 작성합니다.

				Gender	
				Male	Female
Get news from internet	No	Age of respondent	Mean	47.0	48.8
			Median	45.0	46.0
		Highest year of school completed	Mean	13.4	13.1
			Median	13.0	12.0
		Hours per day watching TV	Mean	3.2	3.4
			Median	2.0	3.0
	Yes	Age of respondent	Missing	213	282
			Mean	38.7	41.1
			Median	35.0	38.0
		Highest year of school completed	Mean	13.2	13.3
			Median	13.0	13.0
		Hours per day watching TV	Mean	2.1	2.1
			Median	2.0	2.0
			Missing	0	0

그림 45. 범주형 행 및 열 변수에 따라 그룹화한 척도 요약

척도변수 내에 범주형 변수 중첩

위의 테이블은 필요한 정보를 제공하기는 하지만 해석하기 쉬운 형식은 아닙니다. 예를 들어, 인터넷을 통해 뉴스를 보는 사람과 그렇지 않은 사람의 평균 연령을 비교할 때는 각각의 평균값이 테이블에서 서로 떨어져 있는 경우보다 나란히 위치해 있는 경우 훨씬 비교하기 쉽습니다. 두 행 변수의 위치를 바꾸어 세 개의 척도변수 내에 범주형 그룹화 변수를 중첩하면 표모양이 훨씬 향상됩니다. 척도변수의 경우 중첩 수준이 통계 소스 변수에 영향을 주지 않습니다. 척도변수는 중첩 수준에 상관 없이 항상 통계 소스 변수가 됩니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 반응자연령을 클릭하고 Ctrl 키를 누른 채 피교육연수, 하루 TV시청 시간을 차례로 클릭하여 세 개의 척도변수를 모두 선택합니다.
3. 세 개의 척도변수를 가장 왼쪽 행 영역으로 끌어 인터넷으로부터 뉴스얻음 범주형 변수를 세 개의 척도변수 각각에 중첩시킵니다.
4. 확인을 클릭하여 테이블을 작성합니다.

				Gender	
				Male	Female
Age of respondent	Get news from internet	No	Mean	47.0	48.8
			Median	45.0	46.0
		Yes	Mean	38.7	41.1
			Median	35.0	38.0
Highest year of school completed	Get news from internet	No	Mean	13.4	13.1
			Median	13.0	12.0
		Yes	Mean	13.2	13.3
			Median	13.0	13.0
Hours per day watching TV	Get news from internet	No	Mean	3.2	3.4
			Median	2.0	3.0
			Missing	213	282
		Yes	Mean	2.1	2.1
			Median	2.0	2.0
			Missing	0	0

그림 46. 누적 척도변수 내에 중첩된 범주형 행 변수

테이블에서 강조하려는 관계나 비교 사항에 따라 중첩 순서를 다르게 선택할 수 있습니다. 척도변수의 중첩 순서를 변경해도 요약 통계 값은 변경되지 않으며 단지 테이블에서의 상대적 위치만 변경됩니다.

제 9 장 신뢰구간

신뢰구간 및 표준 오류는 다수의 테이블 통계에 사용할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 최고 차수를 캔버스 분할창의 행 영역으로 이동시킵니다.
3. 요약 통계...를 클릭합니다.
4. 요약 통계 대화 상자의 통계 목록에서 개수 옆의 아이콘을 클릭하여 개수 관련 통계 목록을 펼칩니다.
5. 열 N%, 열 개수 %의 하위 CL, 열 개수 %의 상위 CL 및 열 개수 %의 표준 오류 통계를 표시 영역으로 이동시킵니다.
6. 신뢰구간 그룹에서 수준(%)에 99를 입력합니다.
7. 선택한 항목에 적용을 클릭한 다음 닫기를 클릭합니다.
8. 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	99.0% Lower CL for Column N %	99.0% Upper CL for Column N %	Standard Error of Column N %
Highest degree	LT High school	430	15.2%	13.6%	17.0%	0.7%
	High school	1500	53.2%	50.7%	55.6%	0.9%
	Junior college	209	7.4%	6.2%	8.7%	0.5%
	Bachelor	478	16.9%	15.2%	18.8%	0.7%
	Graduate	205	7.3%	6.1%	8.6%	0.5%

그림 47. 개수, 열 퍼센트 및 신뢰구간을 포함하는 테이블

9. 사용자 정의 테이블 대화 상자를 다시 호출하고 요약 통계...를 클릭합니다.
10. 하위 신뢰 한계의 경우, 레이블을 "하위 신뢰 한계(&[신뢰수준])"으로 변경합니다. 문자열 "&[신뢰수준]"은 레이블의 해당 위치에 지정된 신뢰수준 값을 삽입합니다.
11. 선택한 항목에 적용을 클릭한 다음 닫기를 클릭합니다.
12. 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	Lower Confidence Limit (99.0%)	99.0% Upper CL for Column N %	Standard Error of Column N %
Highest degree	LT High school	430	15.2%	13.6%	17.0%	0.7%
	High school	1500	53.2%	50.7%	55.6%	0.9%
	Junior college	209	7.4%	6.2%	8.7%	0.5%
	Bachelor	478	16.9%	15.2%	18.8%	0.7%
	Graduate	205	7.3%	6.1%	8.6%	0.5%

그림 48. 신뢰구간 레이블이 수정된 테이블

제 10 장 검정 통계량

검정 통계량

세 종류의 유의성 검증을 사용하여 행 변수와 열 변수 간의 관계를 연구할 수 있습니다. 이 장에서는 중첩 및 누적의 효과를 중심으로 각 검정의 결과에 대해 설명합니다. 자세한 정보는 29 페이지의 제 3 장 『범주형 변수를 사용한 누적, 중첩 및 레이어』 주제를 참조하십시오.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

독립성 검정(카이제곱)

독립성에 대한 카이제곱 검정은 두 범주형 변수 간의 관계를 확인하는 데 사용됩니다. 예를 들어, 고용상태가 결혼상태와 관련이 있는지 확인할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록의 고용상태를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. 변수 목록에서 결혼상태를 열 영역으로 끌어다 놓습니다.
4. 요약 통계에 대한 위치에서 행을 선택합니다.
5. 고용상태를 선택하고 정의에서 요약 통계를 클릭합니다.
6. 통계 목록에서 열 N %를 선택하여 출력 목록에 추가합니다.
7. 선택한 항목에 적용을 클릭합니다.
8. 테이블 사용자 정의 대화 상자에서 검정 통계량 탭을 클릭합니다.
9. 독립성 검정(카이제곱)을 선택합니다.
10. 확인을 클릭하여 테이블을 작성하고 카이제곱 검정을 수행합니다.

			Marital status				
			Married	Widowed	Divorced	Separated	Never married
Labor force status	Working full time	Count	778	44	295	58	392
		Column %	57.8%	15.5%	66.1%	62.4%	59.1%
	Working part-time	Count	138	20	35	9	102
		Column %	10.3%	7.1%	7.8%	9.7%	15.4%
	Temporarily not working	Count	23	2	9	1	11
		Column %	1.7%	.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	Count	13	3	10	0	32
		Column %	1.0%	1.1%	2.2%	.0%	4.8%
	Retired	Count	168	150	53	6	17
		Column %	12.5%	53.0%	11.9%	6.5%	2.6%
	School	Count	9	1	7	2	60
		Column %	.7%	.4%	1.6%	2.2%	9.0%
	Keeping house	Count	200	55	25	13	35
		Column %	14.9%	19.4%	5.6%	14.0%	5.3%
Other	Count	16	8	12	4	14	
	Column %	1.2%	2.8%	2.7%	4.3%	2.1%	

그림 49. 결혼상태와 고용상태의 관계

이 테이블은 결혼상태와 고용상태의 교차 분석표로 개수와 열 비율이 요약 통계로 표시됩니다. 각 열의 비율 합계는 100%로 계산됩니다. 이러한 두 변수가 서로 관련이 없을 경우 한조의 열에 대해 행별 열 비율이 비슷해야 합니다. 이 테이블에서는 열 비율이 다르다는 것을 알 수 있는데 카이제곱 검정을 통해 이를 자세히 확인할 수 있습니다.

		Marital status
Labor force status	Chi-square	729.242
	df	28
	Sig.	.000*

*. The Chi-square statistic is significant at the 0.05 level.

그림 50. Pearson의 카이제곱 검정

독립성 검정에서는 고용상태와 결혼상태가 서로 관련이 없다고 가정합니다. 즉, 한조의 열에서의 열 비율은 원래 동일하지만 우연 변동으로 인해 약간의 차이가 관측될 것이라고 가정합니다. 카이제곱 통계를 통해 한조의 열에서의 열 비율이 동일한 경우 기대 셀 개수와 관측 셀 개수 간의 전체 차이를 측정합니다. 카이제곱 통계가 클수록 관측 셀 개수와 기대 셀 개수의 차이가 큼니다. 측정 결과 열 비율이 다르고 독립성 가설이 잘못되었으므로 고용상태와 결혼상태가 서로 관련이 있다는 사실이 입증되었습니다.

계산된 카이제곱 통계 값은 729.242입니다. 이 내용이 독립성 가설을 반박할만한 증거가 되는지 확인하기 위해 통계에 대한 유의수준이 계산됩니다. 유의수준은 자유도가 28인 카이제곱 분포의 확률변수가 729.242보다 큰 확률입니다. 이 값이 검정 통계량 톱에 지정한 알파 수준보다 작기 때문에 0.05 수준에서 독립성 가설을 반박할 수 있습니다. 따라서 고용상태와 결혼상태는 실제로 서로 관련이 있습니다.

독립성 검정에서 중첩과 누적의 효과

독립성 검정 규칙은 다음과 같이 가장 인쪽에 있는 각 하위 테이블에 대해 검정이 수행됩니다. 중첩이 검정에 미치는 영향을 확인하려면 이전 예제를 고려하십시오. 단 결혼상태가 성별 수준 내에 중첩됩니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록에서 성별을 결혼상태 위 캔버스 분할창의 열 영역으로 끌어다 놓습니다.
3. 확인을 클릭하여 테이블을 작성합니다.

		Gender	
		Male	Female
Labor force status	Chi-square	246.637	542.589
	df	28	28
	Sig.	.000 ^{a,1,2}	.000 ^{a,1,2}

*. The Chi-square statistic is significant at the 0.05 level.

1. More than 20% of cells in this sub-table have expected cell counts less than 5.
2. The minimum expected cell count in this sub-table is less than one.

그림 51. Pearson의 카이제곱 검정

결혼상태가 성별 수준 내에 중첩된 상태에서 성별 수준별로 한 번씩, 총 두 번의 검정이 수행됩니다. 각 검정 결과에 따른 유의수준을 통해 남자와 여자 모두에 대해 결혼상태와 고용상태 간의 독립성 가설이 잘못 되었음을 알 수 있습니다. 그러나 테이블 아래에 각 셀 중 20% 이상의 기대 개수가 5보다 작고 최소 기대 셀 개수가 1보다 작다는 내용이 표시됩니다. 따라서 이 테이블만으로는 카이제곱 검정 가설을 만족시킬 수 없으므로 이 검정 결과를 완전히 신뢰할 수는 없습니다.

참고: 표보기에서 이 꼬리말이 셀 경계에서 잘려 표시되지 않을 수 있는데 셀 특성 대화 상자에서 셀 맞춤을 변경하면 꼬리말을 표시할 수 있습니다.

누적이 검정에 미치는 영향을 확인하려면 다음과 같이 하십시오.

4. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
5. 변수 목록에서 최종학력을 고용상태 아래 행 영역으로 끌어다 놓습니다.
6. 확인을 클릭하여 테이블을 작성합니다.

		Gender	
		Male	Female
		Marital status	Marital status
Labor force status	Chi-square	246.637	542.589
	df	28	28
	Sig.	.000 ^{*,1,2}	.000 ^{*,1,2}
Highest degree	Chi-square	43.844	105.506
	df	16	16
	Sig.	.000*	.000*

*. The Chi-square statistic is significant at the 0.05 level.

1. More than 20% of cells in this sub-table have expected cell counts less than 5.

2. The minimum expected cell count in this sub-table is less than one.

그림 52. Pearson의 카이제곱 검정

최종학력을 고용상태로 누적인 상태에서 결혼상태와 고용상태 및 결혼상태와 최종학력에 대한 독립성 검정을 성별 수준별로 한 번씩, 총 네 번 수행합니다. 결혼상태와 고용상태에 대한 검정 결과는 앞의 결과와 동일합니다. 결혼상태와 최종학력에 대한 검정 결과에 따르면 이 두 변수가 서로 독립적이지 않음을 알 수 있습니다.

열 평균 비교

열 평균 검정은 열의 범주형 변수와 행의 연속형 변수 간의 관계를 확인하는 데 사용됩니다. 또한, 이 검정 결과를 사용하여 연속성 변수의 평균값을 기준으로 범주형 변수의 상대적 범주 순위를 확인할 수 있습니다. 예를 들어, 하루 TV 시청 시간과 신문으로부터 뉴스얼음이 서로 관련이 있는지 확인할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 재설정을 클릭하여 모든 탭에 대한 기본값 설정을 복원합니다.
3. 테이블 작성기에서 변수 목록의 하루 TV 시청 시간을 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 신문으로부터 뉴스얼음을 열 영역으로 끌어다 놓습니다.
5. 하루 TV 시청 시간을 선택하고 정의에서 요약 통계를 클릭합니다.
6. 형식으로 nnnn을 선택합니다.
7. 표시할 소수점 이하 자릿수로 2를 선택합니다. 그렇게 하면 형식이 nnnn.nn으로 표시됩니다.
8. 선택한 항목에 적용을 클릭합니다.
9. 테이블 사용자 정의 대화 상자에서 검정 통계량 탭을 클릭합니다.
10. 열 평균 비교(T 검정)를 선택합니다.
11. 확인을 클릭하여 테이블을 작성하고 열 평균 검정을 수행합니다.

	Get news from newspapers	
	No	Yes
	Mean	Mean
Hours per day watching TV	2.92	2.74

그림 53. 하루 TV 시청 시간과 신문으로부터 뉴스얻음의 관계

이 테이블에는 신문에서 뉴스를 보는 사람과 그렇지 않는 사람에 대한 하루 TV 시청 시간 평균이 표시됩니다. 이러한 평균의 차이를 관측해 보면 신문에서 뉴스를 보지 않는 사람이 신문에서 뉴스를 보는 사람보다 하루에 TV를 0.18시간 더 시청한다는 사실을 알 수 있습니다. 이 차이가 우연 변동에 의한 것인지를 알아보려면 열 평균 검정을 확인합니다.

	Get news from newspapers	
	No	Yes
	(A)	(B)
Hours per day watching TV		

그림 54. 열 평균 비교

열 평균 검정 테이블에서는 열 변수의 각 범주에 문자 키를 할당합니다. 신문으로부터 뉴스얻음(E)의 경우 아니오 범주에 A, 예 범주에 B가 할당됩니다. t 검정을 사용하여 각 대응 열의 열 평균을 비교합니다. 열이 두 개뿐이므로 검정은 한 번만 수행됩니다. 각 유의 쌍에 대해 평균이 작은 범주의 키가 평균이 큰 범주 아래에 나타납니다. 표 셀에 키에 대한 보고 내용이 없는 것은 열 평균이 통계적으로 차이가 나지 않기 때문입니다.

APA 스타일 표기법의 유의수준 결과

유의수준 결과를 별도의 테이블에 표시하지 않으려면 기본 테이블에 표시되도록 선택할 수 있습니다. 유의수준 결과는 아래첨자가 있는 APA 스타일 표기법을 사용하여 식별됩니다. 이전 단계를 수행하여 열 평균을 비교하 되 검정 통계량 탭에 다음과 같은 변경사항을 수행하십시오.

1. 유의차 식별 영역에서 **APA** 스타일 아래첨자를 사용하는 기본 테이블에서를 선택합니다.
2. 확인을 클릭하여 테이블을 작성하고 APA 스타일 표기법을 사용하는 열 평균 검정을 가져옵니다.

	Get news from newspapers	
	No	Yes
	Mean	Mean
Hours per day watching TV	2.92 _a	2.74 _a

그림 55. APA 스타일 표기법을 사용하여 열 평균 비교

열 평균 검정 테이블은 열 변수의 범주에 아래첨자를 할당합니다. 각 열 쌍의 경우, t 검정을 사용하여 열 평균이 비교됩니다. 값 쌍이 유의적으로 다른 경우, 해당 값에 할당된 아래첨자와 다른 아래첨자가 지정됩니다. 열이 두 개뿐이므로 검정은 한 번만 수행됩니다. 이 예제에서는 열 평균의 아래첨자가 동일하므로 열 평균이 통계적으로 다르지 않습니다.

열 평균 검정에서 중첩과 누적의 효과

열 평균 검정 규칙은 다음과 같이 가장 안쪽에 있는 각 하위 테이블에 대해 대응별 검정이 수행됩니다 중첩이 검정에 미치는 영향을 확인하려면 이전 예제를 고려하십시오. 단 하루 TV 시청 시간이 고용상태 수준 내에 중첩됩니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록의 고용상태를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. 확인을 클릭하여 테이블을 작성합니다.

			Get news from newspapers	
			No	Yes
			(A)	(B)
Labor force status	Working full time	Hours per day watching TV	B	
	Working part-time	Hours per day watching TV		
	Temporarily not working	Hours per day watching TV		
	Unemployed, laid off	Hours per day watching TV		
	Retired	Hours per day watching TV		
	School	Hours per day watching TV		
	Keeping house	Hours per day watching TV		
	Other	Hours per day watching TV		

그림 56. 열 평균 비교

하루 TV 시청 시간이 고용상태 내에 중첩된 상태에서 고용상태 수준별로 한 번씩 총 7회의 열 평균 검정이 수행됩니다. 신문으로부터 뉴스얻음 범주에 동일한 문자 키가 할당됩니다. 전일근무 반응자의 경우 B 키가 A 열에 나타납니다. 이는 상근직 직원의 하루 TV시청시간 평균값이 신문에서 뉴스를 보는 사람에 비해 낮다는 의미입니다. 다른 키는 열에 나타나지 않으므로 열 평균이 통계적으로 큰 차이가 나지 않는다고 결론지을 수 있습니다.

Bonferroni 조정. 검정이 여러 번 수행되는 경우 열 평균 검정에 Bonferroni 조정을 적용하여 검정 통계량 탭에서 지정한 알파 수준 또는 양의 False 비율을 각 검정군에 적용해야 합니다. 이 테이블에서는 7회의 검정이 수행되더라도 검정군별로 대응 열이 하나만 비교하므로 Bonferroni 조정이 적용되지 않았습니다.

누적이 검정에 미치는 영향을 확인하려면 다음과 같이 하십시오.

4. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
5. 변수 목록에서 인터넷으로부터 뉴스얻음을 신문으로부터 뉴스얻음 왼쪽의 열 영역으로 끌어다 놓습니다.
6. 확인을 클릭하여 테이블을 작성합니다.

			Get news from internet		Get news from newspapers	
			No	Yes	No	Yes
			(A)	(B)	(A)	(B)
Labor force status	Working full time	Hours per day watching TV	B		B	
	Working part-time	Hours per day watching TV	B			
	Temporarily not working	Hours per day watching TV				
	Unemployed, laid off	Hours per day watching TV	B			
	Retired	Hours per day watching TV	B			
	School	Hours per day watching TV	B			
	Keeping house	Hours per day watching TV	B			
	Other	Hours per day watching TV	B			

그림 57. 열 평균 비교

인터넷으로부터 뉴스얻음을 신문으로부터 뉴스얻음으로 누적한 상태에서 인터넷으로부터 뉴스얻음과 신문으로부터 뉴스얻음에 대한 열 평균 검정을 고용상태 수준별로 한 번씩, 총 14회 수행합니다. 또한, 검정군별로 대응 열이 하나만 비교되므로 Bonferroni 조정이 적용되지 않습니다. 신문으로부터 뉴스얻음에 대한 검정은 앞의 결과와 동일합니다. 인터넷으로부터 뉴스얻음(E)의 경우 아니오 범주에 A, 예 범주에 B가 할당됩니다. 일시적 휴직 반응자를 제외하고 열 평균 검정군에 대해 A 열에 B 키가 보고됩니다. 이는 인터넷에서 뉴스를 보는 사람의 하루 TV시청시간 평균값이 신문에서 뉴스를 보지 않는 사람보다 낮다는 의미입니다. 일시적 휴직에 대해서는 키가 보고되지 않으므로 해당 반응자의 열 평균이 통계적으로 차이가 나지 않습니다.

열 비율 비교

열 비율 검정은 행 범주형 변수의 범주 비율을 기준으로 열 범주형 변수의 상대적 범주 순위를 결정하는 데 사용됩니다. 예를 들어, 카이제곱 검정을 사용하여 고용상태와 결혼상태가 독립적이지 않음을 확인한 후 이 관계에 영향을 주는 행과 열을 식별할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 재설정을 클릭하여 모든 탭에 대한 기본값 설정을 복원합니다.
3. 테이블 작성기에서 변수 목록의 고용상태를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 결혼상태를 열 영역으로 끌어다 놓습니다.
5. 고용상태를 선택하고 정의에서 요약 통계를 클릭합니다.
6. 통계 목록에서 열 N %를 선택하여 출력 목록에 추가합니다.
7. 출력 목록에서 개수를 선택 해제합니다.

8. 선택한 항목에 적용을 클릭합니다.
9. 테이블 사용자 정의 대화 상자에서 검정 통계량 탭을 클릭합니다.
10. 열 비율 비교(Z 검정)를 선택합니다.
11. 확인을 클릭하여 테이블을 작성하고 열 비율 검정을 수행합니다.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		Column %	Column %	Column %	Column %	Column %
Labor force status	Working full time	57.8%	15.5%	66.1%	62.4%	59.1%
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4%
	Temporarily not working	1.7%	.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	.0%	4.8%
	Retired	12.5%	53.0%	11.9%	6.5%	2.6%
	School	.7%	.4%	1.6%	2.2%	9.0%
	Keeping house	14.9%	19.4%	5.6%	14.0%	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

그림 58. 결혼상태와 고용상태의 관계

이 테이블은 결혼상태와 고용상태의 교차 분석표로 열 비율이 요약 통계로 표시됩니다.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(A)	(B)	(C)	(D)	(E)
Labor force status	Working full time	B		A B	B	B
	Working part-time					A B C
	Temporarily not working					
	Unemployed, laid off					A B
	Retired	E	A C D E	E		
	School					A B C
	Keeping house	C E	C E		C E	
	Other					

그림 59. 열 비율 비교

열 비율 검정 테이블에서는 열 변수의 각 범주에 문자 키를 할당합니다. 결혼상태의 경우 기혼 범주에 A, 사별 범주에 B 등으로 미혼 범주까지 할당됩니다. Z 검정을 사용하여 각 대응 열에 대해 열 비율을 비교합니다. 고용상태 수준별로 한 번씩, 총 7회의 열 비율 검정이 수행됩니다. 5개의 결혼상태 수준이 있으므로 검정별 $(5*4)/2 = 10$ 개의 대응 열이 비교되고 Bonferroni 조정을 사용하여 유의수준을 조정합니다. 각 유의 쌍에 대한 비율이 작은 범주의 키가 비율이 큰 범위 아래에 나타납니다.

전일근무와 관련된 검정의 경우 B 키가 다른 범주의 열에 각각 나타나며 A 키는 C 범주의 열에만 나타납니다. 다른 열에는 다른 키가 보고되지 않습니다. 따라서 이혼한 상근직 직원의 비율이 기혼 상근직 직원 비율보

다 높고, 기혼 상근직 직원의 비율이 사별한 상근직 직원의 비율보다 높음을 알 수 있습니다. 별거 중이거나 미혼인 상근직 직원의 비율은 이혼했거나 기혼인 상근직 직원의 비율과 구분할 수는 없지만 사별한 상근직 직원의 비율보다는 높습니다.

파트타임 근무 또는 학생과 관련된 검정의 경우 A, B, C 키가 E 범주 열에만 나타납니다. 다른 열에는 다른 키가 보고되지 않습니다. 따라서 미혼인 학생 또는 미혼인 파트타임 근무자의 비율이 학생이거나 파트타임으로 근무 중인 기혼, 사별한 또는 이혼한 사람의 비율보다 높습니다.

일시적 휴직 또는 기타 노동 상태와 관련된 검정의 경우 모든 열에 키가 보고되지 않습니다. 따라서 일시적 실업 상태이거나 범주화되지 않은 고용 상태에 속하는 기혼, 사별, 이혼, 별거 또는 미혼인 사람의 차이는 식별할 수 없습니다.

은퇴와 관련된 검정에서는 퇴직한 사별 직원의 비율이 다른 모든 결혼 여부 범주에 속하는 퇴직한 사람의 비율보다 높습니다. 또한 퇴직자 중 기혼이거나 이혼한 사람의 비율이 미혼인 사람의 비율보다 높습니다.

기혼, 사별 또는 별거 상태이고 주택을 소유한 사람의 비율이 이혼 또는 미혼 상태이고 주택을 소유한 사람의 비율보다 높습니다.

미취업, 해고 상태이고 미혼인 사람의 비율이 미취업 상태이고 기혼 또는 사별인 사람의 비율보다 높습니다. 또한, 별거 열에는 "."가 표시되어 있습니다. 이 표시는 미취업 행에서 별거 중인 사람에 대한 관측 비율이 0 또는 1이기 때문에 실직 상태인 반응자에 대해 해당 열을 사용하여 비교할 수 없음을 나타냅니다.

기본 테이블에 유의성 결과 병합

유의수준 결과를 별도의 테이블에 표시하지 않으려면 기본 테이블에 표시되도록 선택할 수 있습니다. 이전 단계를 수행하여 열 비율을 비교하되 검정 통계량 탭에 다음과 같은 변경사항을 수행하십시오.

1. 유의차 식별 영역에서 기본 테이블에서 선택합니다.
2. 확인을 클릭하여 테이블을 작성합니다.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(a)	(b)	(c)	(d)	(e)
		Column N %	Column N %	Column N %	Column N %	Column N %
Labor force status	Working full time	57.8% b	15.5%	66.1% a b	62.4% b	59.1% b
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4% a b c
	Temporarily not working	1.7%	0.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	0.0% ¹	4.8% a b
	Retired	12.5% e	53.0% a c d e	11.9% e	6.5%	2.6%
	School	0.7%	0.4%	1.6%	2.2%	9.0% a b c
	Keeping house	14.9% c e	19.4% c e	5.6%	14.0% c e	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

그림 60. 기본 테이블에 병합된 유의성 결과

3. 뷰어 창에서 테이블을 두 번 클릭하여 활성화합니다.
4. 문자 키가 포함된 열 레이블 셀 중 하나에 마우스를 놓으십시오. 예를 들어, "(a)" 레이블이 있는 셀 위에 마우스를 놓으십시오.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(a)	(b)	(c)	(d)	(e)
		Column N %	Column N %	Column N %	Column N %	Column N %
Labor force status	Working full time	57.8% b	15.5%	66.1% a b	62.4% b	59.1% b
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4% a b c
	Temporarily not working	1.7%	0.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	0.0% ¹	4.8% a b
	Retired	12.5% e	53.0% a c d e	11.9% e	6.5%	2.6%
	School	0.7%	0.4%	1.6%	2.2%	9.0% a b c
	Keeping house	14.9% c e	19.4% c e	5.6%	14.0% c e	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

그림 61. 선택한 문자 키가 포함된 모든 셀 강조표시

해당 키를 포함하는 모든 셀이 테이블에서 강조표시됩니다.

5. 해당 문자 키가 있는 열 레이블 셀을 마우스 오른쪽 단추로 클릭합니다. 컨텍스트 메뉴에서 선택 > 이 유형의 키가 있는 모든 셀 선택을 선택합니다.

		Marital status				
		Married	Widowed	Divorced	Separated	Never married
		(a)	(b)	(c)	(d)	(e)
		Column N %	Column N %	Column N %	Column N %	Column N %
Labor force status	Working full time	57.8% b	15.5%	66.1% a b	62.4% b	59.1% b
	Working part-time	10.3%	7.1%	7.8%	9.7%	15.4% a b c
	Temporarily not working	1.7%	0.7%	2.0%	1.1%	1.7%
	Unemployed, laid off	1.0%	1.1%	2.2%	0.0% ¹	4.8% a b
	Retired	12.5% e	53.0% a c d e	11.9% e	6.5%	2.6%
	School	0.7%	0.4%	1.6%	2.2%	9.0% a b c
	Keeping house	14.9% c e	19.4% c e	5.6%	14.0% c e	5.3%
	Other	1.2%	2.8%	2.7%	4.3%	2.1%

그림 62. 선택한 문자 키가 포함된 모든 셀 선택

해당 키를 포함하는 모든 셀이 테이블에서 선택됩니다.

열 비율 검정에서 중첩과 누적의 효과

열 비율 검정 규칙은 다음과 같이 가장 안쪽에 있는 각 하위 테이블에 대해 대응별 검정이 수행됩니다 중첩이 검정에 미치는 영향을 확인하려면 이전 예제를 고려하십시오. 단 고용상태가 성별 수준 내에 중첩됩니다.

1. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록의 성별을 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. 확인을 클릭하여 테이블을 작성합니다.

				Marital status				
				Married	Widowed	Divorced	Separated	Never married
				(A)	(B)	(C)	(D)	(E)
Gender	Male	Labor force status	Working full time	B		B	B	B
			Working part-time					A
			Temporarily not working					
			Unemployed, laid off					A
			Retired	E	A C D E	E		
			School					A C
			Keeping house					
	Female	Labor force status	Working full time	B		A B	B	B
			Working part-time	B				B
			Temporarily not working					
			Unemployed, laid off					A
			Retired	E	A C D E	E		
			School					A B C
			Keeping house	C E	C E		C	
		Other						

그림 63. 열 비율 비교

고용상태가 성별 수준 내에 중첩된 상태로 고용상태 수준에 대한 열 비율 검정이 성별 수준별로 한 번씩, 총 14회 수행됩니다. 결혼상태 범주에 앞부분의 예제에서와 동일한 방식으로 문자 키가 할당됩니다.

테이블 결과에 대해 몇 가지 알아야 할 사항이 있습니다.

- 검정 횟수가 많아질수록 열 비율이 0인 열이 많아집니다. 별거 중인 반응자와 사별한 남자의 비율은 거의 일치합니다.
- 앞부분에서 가사일 전담 상태인 반응자에 대해 살펴본 열 차이는 여자에 해당합니다.

누적이 검정에 미치는 영향을 확인하려면 다음과 같이 하십시오.

4. 테이블 작성기를 다시 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
5. 변수 목록에서 최종학력을 성별 아래 행 영역으로 끌어다 놓습니다.
6. 확인을 클릭하여 테이블을 작성합니다.

				Marital status				
				Married	Widowed	Divorced	Separated	Never married
				(A)	(B)	(C)	(D)	(E)
Gender	Male	Labor force status	Working full time	B		B	B	B
			Working part-time					A
			Temporarily not working					
			Unemployed, laid off					A
			Retired	E	A C D E	E		
			School					A C
			Keeping house					
	Female	Labor force status	Working full time	B		A B	B	B
			Working part-time	B				B
			Temporarily not working					
			Unemployed, laid off					A
			Retired	E	A C D E	E		
			School					A B C
			Keeping house	C E	C E		C	
Highest degree		Other						
			LT High school		A C E			
			High school					
			Junior college	B		B		B
			Bachelor	B				B
Graduate	B							

그림 64. 열 비율 비교

최종학력을 성별로 누적한 상태에서 앞에서 설명한 14회의 검정에 최종학력 수준별로 한 번씩의 검정(총 5회)을 추가하여 총 19회의 열 평균 검정이 수행됩니다. 결혼상태 범주에 앞부분의 예제에서와 동일한 방식으로 문자 키가 할당됩니다.

테이블 결과에 대해 몇 가지 알아야 할 사항이 있습니다.

- 14회의 검정 결과는 앞부분의 예제에서 실행해 본 결과와 동일합니다.
- 고졸 미만 학력을 가진 사별한 사람의 비율은 동일 학력의 기혼, 이혼 또는 미혼 반응자 비율보다 높습니다.
- 고졸 학력 이상에서는 사별보다 기혼, 이혼, 미혼의 비율이 높습니다.

가중치와 다중 응답 세트에 대한 참고 사항

케이스 가중치는 변수 중의 하나가 다중반응변수인 경우에도 언제나 반응이 아닌 개수를 기준으로 합니다.

제 11 장 다중 응답 세트

테이블 사용자 정의 및 도표 작성기에서는 다중 응답 세트라는 특수 "변수"를 지원합니다. 다중 응답 세트는 실제로 일반적인 의미의 "변수"는 아닙니다. 다중 응답 세트는 데이터 편집기에서 볼 수 없으며 다른 프로시저에서 이 변수를 인식할 수도 없습니다. 다중 응답 세트는 반응자가 두 개 이상의 대답을 선택할 수 있는 질문에 대해 다중 변수를 사용하여 반응을 기록합니다. 다중 응답 세트는 범주형 변수처럼 처리되며 범주형 변수로 수행할 수 있는 대부분의 작업은 다중 응답 세트로도 수행할 수 있습니다.

다중 응답 세트는 데이터 파일의 다중 변수로 구성됩니다. 이 다중 응답 세트는 데이터 파일에서 매우 특이하게 구성되어 있습니다. IBM® SPSS® Statistics 데이터 파일에서 다중 응답 세트를 정의하고 저장할 수는 있지만 다중 응답 세트를 다른 파일 형식에서 가져오거나 다른 파일 형식으로 내보낼 수 없습니다. 데이터 편집기 창의 데이터 메뉴에서 데이터 특성 복사를 사용하면 다른 IBM SPSS Statistics 데이터 파일의 다중 응답 세트를 복사할 수 있습니다.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 지정합니다.

개수, 반응, 퍼센트 및 총계

척도변수에서 사용할 수 있는 모든 요약 통계는 다중 응답 세트에서도 사용할 수 있습니다. 몇 가지 요약 통계를 다중 응답 세트에도 사용할 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 변수 목록의 새 소스 (이것은 다중 응답 세트 *\$mltnews*에 대한 설명 레이블입니다)를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

변수 목록의 "변수" 옆에 있는 아이콘은 이것을 다중 이분 세트로 식별합니다.



그림 65. 다중 이분 세트 아이콘

다중 이분 세트의 경우 각 "범주"는 사실상 개별 변수이며 범주 레이블은 변수 레이블입니다(또는 정의된 변수 레이블이 없는 변수에 대한 변수 이름입니다). 이 예에서 표시될 개수는 그룹의 각 변수에 대한 예 반응이 포함된 케이스 수를 표현합니다.

3. 캔버스 분할창의 테이블 미리보기에서 새 소스를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
4. 범주 및 총계 대화 상자에서 총계 확인란을 선택한 다음 적용을 클릭합니다.
5. 새 소스를 다시 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
6. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
7. 선택한 항목에 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %
News sources	Get news from internet	867	41.7%
	Get news from radio	551	26.5%
	Get news from television	1077	51.8%
	Get news from news magazines	294	14.1%
	Get news from newspapers	805	38.7%
	Total	2081	100.0%

그림 66. 다중 이분형 개수 및 열 퍼센트

추가되지 않는 총계

테이블의 숫자를 살펴보면 "총계"와 합산되는 중으로 생각되는 값 사이에 매우 큰 오차가 존재하는 것을 알 수 있습니다. 특히 총계는 생각보다 훨씬 낮게 표시됩니다. 이것은 테이블의 각 "범주"에 대한 개수가 그 변수에 대해 값 1(예 반응)을 갖는 케이스 수이기 때문입니다. 또 다중 이분 세트의 5개의 모든 변수에 대한 예 반응의 총계는 데이터 파일의 총 케이스 개수를 쉽게 초과할 수 있을 것이기 때문입니다.

그러나 총 "개수"는 그룹에서 최소한 하나의 변수에 대해 예 반응을 갖는 케이스의 총계입니다. 이 총 "개수"는 데이터 파일의 총 케이스 개수를 절대 초과할 수 없습니다. 이 예에서 2,081이란 총계는 데이터 파일의 총 케이스 개수보다 거의 800개나 적은 수치입니다. 이 변수중 어느 것도 결측값을 갖고 있지 않다면 이것은 거의 800명의 설문조사 반응자가 그 소스에서 어느 것으로부터도 뉴스를 가져오지 않았다는 것을 의미합니다. 총계는 열 퍼센트의 기준이므로 이 예에서 열 퍼센트의 총계는 총 열 퍼센트를 100% 표시한 것보다 더 많습니다.

추가되는 총계

"개수"라는 말은 보통 매우 명백한 용어이지만 위의 예는 다중 응답 세트에서 총계를 낼 때와 같은 상황에서 이 용어가 얼마나 혼동될 수 있는지를 잘 증명해 줍니다. 이 개수에서는 반응이 종종 여러분이 정말로 원하는 요약 통계가 되기 때문입니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 새 소스를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
3. 요약 통계 대화 상자에서 통계 목록의 반응을 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.

4. 통계 목록에서 열 반응 %를 선택하고 화살표 키를 눌러 출력 목록에 추가합니다.
5. 선택한 항목에 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	Responses	Column Responses %	
News sources	Get news from internet	867	41.7%	867	24.1%	24.1%
	Get news from radio	551	26.5%	551	15.3%	15.3%
	Get news from television	1077	51.8%	1077	30.0%	30.0%
	Get news from news magazines	294	14.1%	294	8.2%	8.2%
	Get news from newspapers	805	38.7%	805	22.4%	22.4%
Total		2081	100.0%	3594	100.0%	100.0%

그림 67. 다중 이분형 반응 및 열 반응 퍼센트

다중 이분 세트 테이블에서 각 "범주"의 경우 반응은 개수와 동일하며 다중 이분 세트의 경우는 항상 그렇습니다. 그러나 총계는 매우 다릅니다. 반응의 총계는 3,594입니다. 총계보다 1,500개 많으며 데이터 파일의 총 케이스 개수보다 700개가 더 많습니다.

퍼센트의 경우 열 N % 및 열 반응 %의 총계는 모두 100%이지만 다중 이분 세트에서 각 범주에 대한 퍼센트는 열 반응 퍼센트보다 훨씬 낮습니다. 이것은 열 반응 퍼센트의 퍼센트 기준이 반응의 총 개수이기 때문입니다. 총 개수는 이 경우에는 3,594이며 열 퍼센트 기준인 2,081보다 훨씬 낮은 퍼센트입니다.

100%보다 큰 퍼센트 총계.

열 퍼센트와 열 반응 퍼센트 모두 100%의 전체 퍼센트를 산출할지라도 이 예에서 열 N % 열의 개별 값을 합하면 분명 100%보다 큼니다. 그러면 만일 총 반응이 아닌 총계를 기준으로 퍼센트를 표시하면서 "총" 퍼센트가 개별 범주 퍼센트의 총계를 정확하게 반영하도록 하려면 어떻게 합니까?

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 새 소스를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
3. 요약 통계 대화 상자에서 통계 목록의 열 응답 %(기준: 개수)를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
4. 선택한 항목에 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	Responses	Column Responses %	Column Responses % (Base: Count)
News sources	Get news from internet	867	41.7%	867	24.1%	41.7%
	Get news from radio	551	26.5%	551	15.3%	26.5%
	Get news from television	1077	51.8%	1077	30.0%	51.8%
	Get news from news magazines	294	14.1%	294	8.2%	14.1%
	Get news from newspapers	805	38.7%	805	22.4%	38.7%
Total		2081	100.0%	3594	100.0%	172.7%

그림 68. 퍼센트 기준이 개수인 열 반응 퍼센트

다른 변수와 함께 다중 응답 세트 사용

대개 범주형 변수와 마찬가지로 다중 응답 세트를 사용할 수 있습니다. 예를 들어 범주형 변수를 포함한 다중 응답 세트에 관한 교차 분석표를 작성하거나 범주형 변수 내에 다중 응답 세트를 중첩시킬 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록에서 미리보기 분할창의 행 영역의 왼쪽 사이드로 성별을 끌어놓아 성별 범주 안에 다중 응답 세트 새 소스를 중첩시킵니다.
3. 캔버스 분할창의 테이블 미리보기에서 성별을 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 변수 레이블 표시를 선택 취소합니다.
4. 새 소스에도 같은 작업을 수행합니다.

이렇게 하면 테이블에서 변수 레이블이 있는 열을 제거하게 됩니다(이 경우에는 변수 레이블이 있는 열이 실제로 필요하지 않으므로).

5. 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	Responses	Column Responses %	Column Responses % (Base: Count)
Male	Get news from internet	359	40.1%	359	23.3%	40.1%
	Get news from radio	233	26.0%	233	15.1%	26.0%
	Get news from television	451	50.3%	451	29.3%	50.3%
	Get news from news magazines	121	13.5%	121	7.9%	13.5%
	Get news from newspapers	375	41.9%	375	24.4%	41.9%
	Total	896	100.0%	1539	100.0%	171.8%
Female	Get news from internet	508	42.9%	508	24.7%	42.9%
	Get news from radio	318	26.8%	318	15.5%	26.8%
	Get news from television	626	52.8%	626	30.5%	52.8%
	Get news from news magazines	173	14.6%	173	8.4%	14.6%
	Get news from newspapers	430	36.3%	430	20.9%	36.3%
	Total	1185	100.0%	2055	100.0%	173.4%

그림 69. 범주형 변수 안에 중첩된 다중 응답 세트

통계량 소스 변수 및 사용 가능한 요약 통계

테이블에 척도변수가 없는 경우 범주형 변수 및 다중 응답 세트를 통계 소스 변수에 관하여 동일한 방법으로 취급합니다. 통계 소스 차원에서 가장 인쪽에 있는 중첩 변수는 통계 소스 변수입니다. 다중 응답 세트에만 할당될 수 있는 요약 통계가 몇 가지 있기 때문에 이것은 특별한 다중반응 요약 통계 중 어느 하나를 원할 경우 다중 응답 세트가 통계 소스 차원에서 가장 인쪽에 중첩된 변수가 되어야 함을 의미합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 새 소스를 성별 왼쪽으로 끌어 놓아 중첩 순서를 변경합니다.

반응, 열 반응 퍼센트 등 특수한 다중반응 요약 통계는 모두 테이블 미리보기에서 제거됩니다. 왜냐하면 범주형 변수인 성별이 이제 가장 인쪽에 중첩된 변수이며 따라서 통계 소스 변수가 되기 때문입니다.

다행히 테이블 작성기는 이 설정을 "기억합니다". 새 소스를 성별 안에 중첩된 이전의 위치로 다시 이동할 경우 반응과 관련된 모든 요약 통계가 테이블 미리보기로 복원됩니다.

다중 범주 세트 및 중복 응답

다중 범주 세트는 다중 이분 세트에 없는 한 가지 특징인 중복 응답을 셀 수 있는 기능을 제공합니다. 수많은 경우에 다중 범주 세트의 중복 반응은 코딩 오류를 나타낼 경우가 있습니다. 예를 들면 "최고의 자동차를 제작한다고 생각되는 세 나라는 어느 나라입니까?"와 같은 설문조사의 경우 스웨덴, 독일 및 스웨덴이라는 반응은 유효하지 않은 반응일 것입니다.

그러나 중복 반응은 완벽하게 유효한 다른 경우도 있을 수 있습니다. 예를 들어, 질문이 "귀하가 마지막으로 구입한 자동차 세 대는 어느 나라에서 제조된 것입니까?"일 경우 스웨덴, 독일 및 스웨덴이라는 반응이 납득할 만한 반응일 것입니다.

테이블 사용자 정의는 다중 범주 세트에서 중복 반응의 기회를 제공합니다. 기본적으로 중복 반응은 세지 않으나 중복 반응을 포함시키도록 요청할 수는 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 이전 설정을 지우려면 재설정을 클릭합니다.
3. 변수 목록의 자동차 제조업체, 가장 최근 차를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.

변수 목록의 "변수" 옆에 있는 아이콘은 이것을 다중 범주 세트로 식별합니다.



그림 70. 다중 범주 세트 아이콘

다중 범주 세트의 경우 표시된 범주는 그룹의 모든 변수에 대하여 공통적인 정의된 값 레이블 세트를 나타냅니다(반면, 다중 이분 세트의 경우 "범주"는 실제로 그룹의 각 변수에 대한 변수 레이블입니다).

4. 캔버스 분할창의 테이블 미리보기에서 자동차 제조업체, 가장 최근 차를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
5. 범주 및 총계 대화 상자에서 총계 확인란을 선택한 다음 적용을 클릭합니다.
6. 자동차 제조업체, 가장 최근 차를 다시 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
7. 요약 통계 대화 상자에서 통계 목록의 반응을 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
8. 선택한 항목에 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Responses
Car maker, most recent cars	American	1938	1938
	Japanese	1327	1327
	Korean	695	695
	German	693	693
	Swedish	360	360
	Other	343	343
	Total	2832	5356

그림 71. 다중 범주 세트: 중복이 없는 개수 및 반응

기본적으로, 중복 반응은 세지 않습니다. 따라서 이 테이블에서 개수 및 반응의 각 범주에 대한 값이 동일합니다. 총계만 다릅니다.

9. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
10. 옵션 탭을 클릭합니다.
11. 다중 범주 세트의 중복 반응수 계산을 선택합니다.
12. 확인을 클릭하여 테이블을 작성합니다.

		Count	Responses
Car maker, most recent cars	American	1938	2797
	Japanese	1327	1717
	Korean	695	760
	German	693	754
	Swedish	360	383
	Other	343	359
	Total	2832	6770

그림 72. 중복 반응을 포함한 다중 범주 세트

이 테이블에서 개수 및 반응 열의 값 사이에는 현저한 차이가 존재합니다. 특히 미제 자동차의 경우 수많은 반응자들이 여러 대의 미제 자동차를 소유한 것으로 나타납니다.

다중 응답 세트에 대한 유의성 검증

유의성 검증에서는 범주형 변수를 사용하는 것과 본질상 동일한 방법으로 다중 응답 세트를 사용할 수 있습니다.

- 독립성 검정(카이제곱) 또는 열 비율 비교(z-검정)의 경우 개수에 대해 검정을 수행하며 개수는 테이블에 표시된 요약 통계중 하나가 되어야 합니다.
- 다중 범주 세트의 경우 옵션 탭에서 다중 범주 세트의 중복 반응수 계산을 선택한 경우 열 비율 비교 또는 열 평균 비교(t-검정)를 수행합니다. 자세한 정보는 17 페이지의 『테이블 사용자 정의: 옵션 탭』 주제를 참조하십시오.

다중 응답 세트에 대한 독립성 검정

이 예는 범주형 변수와 다중 응답 세트의 교차 분석표를 생성하며 교차 분석표에서 카이제곱 독립성 검정을 수행합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).

2. 이전 설정을 지우려면 재설정을 클릭합니다.
3. 변수 목록의 새 소스 (이것은 다중 이분 세트 *\$mltnews*에 대한 설명 레이블입니다)를 캔버스 분할창의 열 영역으로 끌어다 놓습니다.
4. 변수 목록의 성별을 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
5. 검정 통계량 탭을 클릭합니다.
6. 독립성 검정(카이제곱)을 선택합니다.
7. 검정에 다중반응변수 포함이 선택되어 있지 않으면 이를 선택합니다.
8. 확인을 클릭하여 프로시저를 실행합니다.

		News sources
Gender	Chi-square	10.266
	df	5
	Sig.	.068

그림 73. 카이제곱 검정 결과

카이제곱 검정에 대한 유의 수준 레벨 0.068은 남성과 여성이 새 소스를 선택함에 있어 별반 다르지 않음을 나타냅니다(통계 유의수준을 판단하기 위한 기준으로 0.05 이하의 유의수준을 사용한다는 것을 가정하에).

다중 응답 세트에 대한 열 평균 비교

이 예는 다중 응답 세트에 의하여 정의된 범주 내에서 척도변수 평균을 계산하며 각 범주 평균을 기타 모든 범주 평균과 유의차가 있는지 비교합니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 이전 설정을 지우려면 재설정을 클릭합니다.
3. 변수 목록의 새 소스 (이것은 다중 이분 세트 *\$mltnews*에 대한 설명 레이블입니다)를 캔버스 분할창의 열 영역으로 끌어다 놓습니다.
4. 반응자 연령을 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
5. 검정 통계량 탭을 클릭합니다.
6. 열 평균 비교(T 검정)를 선택합니다.
7. 검정에 다중반응변수 포함이 선택되어 있지 않으면 이를 선택합니다.
8. 확인을 클릭하여 프로시저를 실행합니다.

	News sources				
	Get news from newspapers	Get news from news magazines	Get news from television	Get news from radio	Get news from internet
	Mean	Mean	Mean	Mean	Mean
Age of respondent	52	40	48	40	40

Comparisons of Column Means

	News sources				
	Get news from newspapers	Get news from news magazines	Get news from television	Get news from radio	Get news from internet
	(A)	(B)	(C)	(D)	(E)
Age of respondent	B C D E		B D E		

Results are based on two-sided tests assuming equal variances with significance level 0.05. For each significant pair, the key of the smaller category appears under the category with larger mean.

그림 74. 유의성 검증 결과

- 다중 응답 세트의 각 범주는 문자(A, B, C, D, E)로 식별되며, 각 범주에 대해서 또 다른 범주의 평균이 그 범주의 평균보다 낮고 큰 차이가 날 경우에는 낮은 평균을 갖는 범주임을 나타내는 문자가 표시됩니다.
- 신문으로부터 뉴스 얻음(A)은 평균 연령이 가장 높으며 다른 모든 범주의 평균은 최고 평균 연령과 큰 차이가 납니다.
- TV로부터 뉴스 얻음(A)은 평균 연령이 두 번째로 높으며 남은 모든 범주의 평균(B, D, E)은 두 번째로 높은 이 평균 연령과 큰 차이가 납니다. (또한 C는 이전에 언급한 대로 A와 큰 차이가 납니다.)
- 잡지로부터 뉴스 얻음, 라디오로부터 뉴스 얻음(D) 및 인터넷으로부터 뉴스 얻음(E)의 평균 연령은 서로 큰 차이가 나지 않습니다.

제 12 장 결측값

많은 데이터 파일에는 일정한 양의 결측 데이터가 들어 있습니다. 이러한 결측 데이터는 여러 가지 요인으로 인해 생성될 수 있습니다. 예를 들어, 설문조사 반응자가 일부 질문에 대답하지 않거나, 특정 변수가 일부 케이스에 적용될 수 없거나, 코딩 오류로 인해 일부 값이 제거될 수 있습니다.

IBM SPSS Statistics에는 다음 두 가지의 결측값이 있습니다.

- **사용자 결측값.** 결측 데이터를 포함하는 것으로 정의된 값입니다. 이러한 값에는 데이터가 결측된 이유를 나타내는 값 레이블(예: 남자가 임신한 경우 해당없음이라는 값 레이블 및 코드 99)이 할당될 수 있습니다.
- **시스템-결측값.** 숫자변수에 값이 없으면 시스템-결측값이 할당됩니다. 이 값은 데이터 편집기의 데이터 보기에서 마침표로 표시됩니다.

여기에는 결측 데이터로 인한 영향을 보완할 수 있을 뿐만 아니라 결측 데이터의 패턴을 분석할 수도 있는 여러 가지 유용한 기능이 제공되지만 그러나 이 장에는 보다 간단한 목표가 있습니다. 테이블 사용자 정의가 결측 데이터를 처리하는 방법과 결측 데이터가 요약 통계 계산에 미치는 영향에 대해 설명합니다.

표본 데이터 파일

이 장의 예제에서는 *missing_values.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오. 변수 한 개와 케이스 10개만 들어 있는 매우 간단한 이 예제 데이터 파일은 결측값의 기본 개념을 설명하도록 만들어졌습니다.

결측값이 없는 테이블

기본적으로 사용자 결측 범주는 테이블 사용자 정의에 표시되지 않습니다. 시스템-결측값은 어떤 경우에도 표시되지 않습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록에 있는 결측값이 포함된 변수(이 파일의 유일한 변수)를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. 캔버스 분할창에서 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
4. 범주 및 총계 대화 상자에서 총계 확인란을 선택한 다음 적용을 클릭합니다.
5. 캔버스 분할창의 테이블 미리보기에서 결측값이 포함된 변수를 마우스 오른쪽 단추로 다시 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
6. 요약 통계 대화 상자에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
7. 선택한 항목에 적용을 클릭합니다.

캔버스 분할창의 테이블 미리보기에 표시된 범주와 테이블 작성기 왼쪽의 변수 목록 아래에 있는 범주 목록에 표시된 범주가 약간 다릅니다. 범주 목록에는 결측값이라는 레이블이 붙은 범주가 있지만 테이블 미리보기에는 결측값 범주가 기본적으로 제외되기 때문에 표시되지 않습니다. 레이블에서 "값"이 여러 개이므로 이 변수에 두 개 이상의 사용자 결측값이 있음을 알 수 있습니다.

8. 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %
Variable with missing values	Low	2	28.6%
	Medium	3	42.9%
	High	2	28.6%
	Total	7	100.0%

그림 75. 결측값이 없는 테이블

이 테이블의 모든 내용은 정확합니다. 범주 값의 정확한 총계가 표시되어 있고, 퍼센트 기준의 총계(예: $3/7 = 0.429$ 또는 42.9%) 값이 정확하게 계산되어 있습니다. 그러나 총계는 데이터 파일에 있는 총 케이스 수가 아니라 결측되지 않은 값이 있는 총 케이스 수 또는 해당 변수에 대해 사용자 결측값이나 시스템-결측값을 포함하지 않는 총 케이스 수입니다.

테이블에 결측값 포함

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 캔버스 분할창의 테이블 미리보기에서 결측값이 포함된 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
3. 범주 및 총계 대화 상자에서 결측값 확인란을 선택한 다음 적용을 클릭합니다.

이제 테이블 미리보기에 결측값 범주가 포함되어 있습니다. 테이블 미리보기에 결측값에 대해 하나의 범주만 표시되어 있지만 테이블에는 모든 사용자 결측 범주가 표시됩니다.

4. 캔버스 분할창의 테이블 미리보기에서 결측값이 포함된 변수를 마우스 오른쪽 단추로 다시 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
5. 요약 통계 대화 상자에서 총계 및 소계에 대한 사용자 정의 요약 통계 확인란을 선택합니다.
6. 사용자 정의 요약 통계 목록에서 유효수를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
7. 합계 N에 대해서도 같은 작업을 수행합니다.
8. 선택한 항목에 적용을 클릭한 다음 테이블 작성기에서 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	Valid N	Total N
Variable with missing values	Low	2	22.2%		
	Medium	3	33.3%		
	High	2	22.2%		
	Don't know	1	11.1%		
	Not applicable	1	11.1%		
	Total	9	100.0%	7	10

그림 76. 결측값이 있는 테이블

잘 모름 및 해당없음이라는 두 개의 정의된 사용자 결측 범주가 테이블에 추가되었습니다. 사용자 결측값이 있는 두 개의 케이스(사용자 결측 범주당 하나씩)가 추가되었기 때문에 총계가 7에서 9로 바뀝니다. 비결측값과 사용자 결측값을 기준으로 하는 열 퍼센트도 달라집니다. 시스템-결측값만 퍼센트 계산에서 제외됩니다.

유효한 N은 결측되지 않은 총 케이스 수(7)를, 합계 N은 사용자 결측값 및 시스템-결측값이 있는 총 케이스 수를 표시합니다. 총 케이스 수는 10으로 개수 열에 총계로 표시된 비결측값과 사용자 결측값의 개수보다 하나가 더 많습니다. 이는 시스템-결측값이 있는 케이스가 하나 더 있기 때문입니다.

9. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
10. 캔버스 분할창의 테이블 미리보기에서 결측값이 포함된 변수를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
11. 총계 및 소계에 대한 사용자 정의 요약 통계 목록이 아닌 위쪽의 통계 목록에서 열-유효수 %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
12. 열 총계 N %에 대해서도 같은 작업을 수행합니다.
13. 이 둘을 총계 및 소계에 대한 사용자 정의 요약 통계 목록에도 추가할 수 있습니다.
14. 선택한 항목에 적용을 클릭한 다음 확인을 클릭하여 테이블을 작성합니다.

		Count	Column N %	Column Valid N %	Column Total N %	Valid N	Total N
Variable with missing values	Low	2	22.2%	28.6%	20.0%		
	Medium	3	33.3%	42.9%	30.0%		
	High	2	22.2%	28.6%	20.0%		
	Don't know	1	11.1%	.0%	10.0%		
	Not applicable	1	11.1%	.0%	10.0%		
	Total	9	100.0%	100.0%	100.0%	7	10

그림 77. 결측값과 유효 퍼센트 및 전체 퍼센트가 있는 테이블

- 열 N %는 비결측값과 테이블에 명백하게 포함되어 있는 사용자 결측값을 기준으로 한 각 범주의 퍼센트입니다.
- 열-유효수 %는 결측값이 없는 유효한 케이스만을 기준으로 한 각 범주의 퍼센트입니다. 이 값들은 사용자 결측값이 포함되지 않은 원래 테이블의 열 퍼센트와 같습니다.
- 열 총계 N %는 사용자 결측값과 시스템-결측값이 있는 케이스를 비롯한 모든 케이스를 기준으로 한 각 범주의 퍼센트입니다. 이 범주의 개별 범주 퍼센트 총계가 90%밖에 되지 않는데 이는 총 10개의 케이스 중

한 개의 케이스(10%)에 시스템-결측값이 포함되어 있기 때문입니다. 시스템-결측값이 있는 케이스가 퍼센트 계산 기준에는 포함되어 있지만 테이블에 해당 범주가 표시되지는 않습니다.

제 13 장 표 형식 지정 및 사용자 정의

표 형식 지정 및 사용자 정의

테이블 사용자 정의를 사용하면 테이블 작성 절차의 일부로 아래와 같은 여러 표 형식 특성을 제어할 수 있습니다.

- 요약 통계에 대한 표시 형식 및 레이블
- 최소 및 최대 데이터 열 너비
- 빈 셀에 표시되는 텍스트 또는 값

이러한 설정을 변경하거나 테이블 작성기 설정을 재설정하거나 다른 데이터 파일을 열기 전까지 테이블 작성기 인터페이스에서 이 설정이 유지되므로 테이블을 작성한 후에 수동으로 편집하지 않고 형식 특성이 같은 테이블을 여러 개 만들 수 있습니다. 테이블 작성기 인터페이스에 있는 붙여넣기 단추를 사용하여 붙여넣기 명령 구문을 구문 창에 붙여넣은 다음 파일로 저장하면 이러한 형식 설정을 다른 모든 테이블 모수와 함께 저장할 수도 있습니다.

또한 피벗 테이블 뷰어에서 사용할 수 있는 모든 형식 지정 공정능력으로 테이블을 작성한 후 여러 가지 표 형식 특성을 바꿀 수도 있습니다. 하지만 이 장에서는 테이블을 작성하기 전에 표 형식 특성을 제어하는 방법을 주로 설명합니다. 피벗 테이블에 대한 자세한 내용을 보려면 도움말 시스템의 색인 탭에 키워드로 피벗 테이블을 입력하십시오.

표본 데이터 파일

이 장의 예제에서는 *survey_sample.sav* 데이터 파일을 사용합니다. 자세한 정보는 데이터 파일 주제를 참조하십시오.

여기에 제공되는 모든 예제는 변수 레이블을 문자순으로 정렬하여 대화 상자에 표시합니다. 변수 목록 출력 특성은 옵션 대화 상자(편집 메뉴, 옵션)의 일반 탭에서 설정합니다.

요약 통계 표시 형식

테이블 사용자 정의에서는 요약 통계에 비교적 다양하게 활용할 수 있는 기본값 형식을 적용하려고 하지만 이러한 기본값을 변경해야 하는 경우도 있을 수 있습니다.

1. 메뉴에서 다음을 선택합니다.

분석 > 테이블 > 테이블 사용자 정의...

2. 테이블 작성기에서 변수 목록에 있는 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
3. TV에 대한 신뢰도를 행 영역의 연령범주 아래로 끌어 놓아 두 개의 변수를 행 차원으로 누적합니다.

4. 캔버스 분할창의 테이블 미리보기에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 모든 행 변수 선택을 선택합니다.
5. 연령범주를 마우스 오른쪽 단추로 다시 클릭하고 팝업 메뉴에서 범주 및 총계를 선택합니다.
6. 범주 및 총계 대화 상자에서 총계를 선택한 다음 적용을 클릭합니다.
7. 캔버스 분할창의 테이블 미리보기에서 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.
8. 통계 목록에서 열 N %를 선택하고 화살표 키를 눌러 출력 목록에 추가합니다.
9. 총계 및 소계에 대한 사용자 정의 요약 통계를 선택합니다.
10. 사용자 정의 요약 통계에서 통계 목록의 열 N %를 선택하고 화살표를 클릭하여 출력 목록에 추가합니다.
11. 평균에 대해 같은 작업을 수행합니다.
12. 그런 다음 모든 항목에 적용을 클릭합니다.

자리표시자 값은 각 요약 통계의 기본값 형식을 반영합니다.

- 개수의 경우 기본값 표시 형식은 소수점 이하 자릿수가 없는 정수인 **nnnn**입니다.
- 퍼센트의 경우 기본값 표시 형식은 소수점 이하 자릿수가 한 자리이며 값 뒤에 퍼센트 기호가 있는 **nnnn.n%**입니다.
- 평균의 경우 두 변수의 기본값 표시 형식은 *다릅니다*.

개수(유효수, 합계 N 등) 또는 퍼센트 형식 이외의 요약 통계의 경우 기본값 표시 형식은 데이터 편집기에서 변수에 대해 정의한 표시 형식입니다. 데이터 편집기의 변수 보기에서 변수를 찾아보면 연령범주(변수 연령범주)는 소수점 이하 자릿수가 2로 정의되어 있는 반면 TV에 대한 신뢰도(변수 TV신뢰도)는 소수점 이하 자릿수가 0으로 정의되어 있습니다.

두 평균값의 소수점 이하 자릿수는 동일하게 표시하는 것이 좋으므로 이러한 경우에는 기본값 형식을 사용하지 않는 것이 좋습니다.

13. 캔버스 분할창의 테이블 미리보기에서 변수 중 하나를 마우스 오른쪽 단추로 클릭하고 팝업 메뉴에서 요약 통계를 선택합니다.

평균의 경우 출력 목록의 형식 셀에 형식이 자동으로 표시되는데, 이는 변수에 대해 정의된 표시 형식이 사용되고 소수점이하자리 셀을 사용할 수 없다는 의미입니다. 소수점 이하 자릿수를 지정하려면 우선 다른 형식을 선택해야 합니다.

14. 사용자 정의 요약 통계 출력 목록에서 평균에 대한 형식 셀을 클릭하고 형식 드롭다운 목록에서 **nnnn**을 선택합니다.
15. 소수점이하자리 셀에 값 1을 입력합니다.
16. 그런 다음 모든 항목에 적용을 클릭하여 두 변수에 이 설정을 적용합니다.

이제 테이블 미리보기에서 두 평균값은 소수점 이하 자릿수가 1인 형식으로 표시됩니다. (지금 바로 이 테이블을 작성할 수도 있지만 실제로 이 변수에 대한 숫자 코드가 1에서 6까지로 제한되기 때문에 연령범주에 대한 "평균"값을 해석하기가 약간 까다롭다는 것을 알 수 있습니다.)

요약 통계에 대한 레이블 표시

요약 통계의 표시 형식을 비롯하여 각 요약 통계에 대한 설명 레이블도 제어할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 테이블 작성기에서 재설정을 클릭하여 이전 설정을 모두 지웁니다.
3. 테이블 작성기에서 변수 목록에 있는 연령범주를 캔버스 분할창의 행 영역으로 끌어다 놓습니다.
4. 변수 목록에서 임금수령방법을 캔버스 분할창의 열 영역으로 끌어다 놓습니다.
5. 캔버스 분할창의 테이블 미리보기에서 연령범주를 마우스 오른쪽 단추로 클릭하고 팝업 컨텍스트 메뉴에서 요약 통계를 선택합니다.
6. 통계 목록에서 열 **N %**를 선택하고 회살표 키를 눌러 출력 목록에 추가합니다.
7. 출력 목록의 레이블 셀에서 열이란 단어를 두 번 클릭하고 셀의 내용을 편집합니다. 레이블에서 열이란 단어를 삭제하여 레이블을 %로 변경합니다.
8. 같은 방법으로 개수에서도 레이블을 *N*으로 바꾸어 레이블 셀을 편집합니다.

이제 열 레이블을 통해 열에 퍼센트가 포함되어 있다는 것을 알 수 있기 때문에 열 **N %** 통계의 형식을 변경하여 불필요한 퍼센트 기호를 제거합니다.

9. 열 *N %*의 형식 셀을 클릭하고 형식의 드롭다운 목록에서 **nnnn.n**을 선택합니다.
10. 그런 다음 선택한 항목에 적용을 클릭합니다.

테이블 미리보기에 수정된 표시 형식과 레이블이 표시됩니다.

11. 확인을 클릭하여 테이블을 작성합니다.

		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		N	%	N	%	N	%	N	%	N	%	N	%
Age category	Less than 25	91	14.0	0	.0	12	9.7	3	2.0	7	3.1	14	7.7
	25 to 34	175	26.9	5	29.4	33	26.6	37	24.8	63	28.0	31	17.1
	35 to 44	185	28.5	5	29.4	42	33.9	45	30.2	66	29.3	61	33.7
	45 to 54	124	19.1	5	29.4	25	20.2	38	25.5	58	25.8	41	22.7
	55 to 64	52	8.0	0	.0	10	8.1	23	15.4	29	12.9	19	10.5
	65 or older	23	3.5	2	11.8	2	1.6	3	2.0	2	.9	15	8.3

그림 78. 요약 통계 레이블이 수정된 테이블

열 너비

이전 예제에 있는 테이블이 약간 넓어 보입니다. 이를 해결하려면 행 변수와 열 변수를 서로 바꾸거나 지나치게 넓어 보이는 열을 좁혀보십시오. (실제로 열을 좁히기 위해 요약 통계 레이블을 줄였습니다.)

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 옵션 탭을 클릭합니다.
3. 데이터 열 너비에서 사용자 정의를 선택합니다.
4. 최대값에 36을 입력합니다. (단위가 포인트로 설정되어 있는지 확인합니다.)
5. 확인을 클릭하여 테이블을 작성합니다.

		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		N	%	N	%	N	%	N	%	N	%	N	%
Age category	Less than 25	91	14.0	0	.0	12	9.7	3	2.0	7	3.1	14	7.7
	25 to 34	175	26.9	5	29.4	33	26.6	37	24.8	63	28.0	31	17.1
	35 to 44	185	28.5	5	29.4	42	33.9	45	30.2	66	29.3	61	33.7
	45 to 54	124	19.1	5	29.4	25	20.2	38	25.5	58	25.8	41	22.7
	55 to 64	52	8.0	0	.0	10	8.1	23	15.4	29	12.9	19	10.5
	65 or older	23	3.5	2	11.8	2	1.6	3	2.0	2	.9	15	8.3

그림 79. 열 너비가 좁혀진 테이블

이제 테이블이 훨씬 간결해졌습니다.

빈 셀의 표시 값

기본적으로 빈 셀(케이스가 없는 셀)에는 0이 표시됩니다. 대신 이 셀에 아무것도 표시하지 않거나(공백으로 둠), 빈 셀에 표시할 텍스트 문자열을 지정할 수 있습니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 옵션 탭을 클릭합니다.
3. 데이터 셀 모양에서 빈 셀에 대해 텍스트를 선택하고 없음을 입력합니다.
4. 확인을 클릭하여 테이블을 작성합니다.

		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		N	%	N	%	N	%	N	%	N	%	N	%
Age category	Less than 25	91	14.0	None	None	12	9.7	3	2.0	7	3.1	14	7.7
	25 to 34	175	26.9	5	29.4	33	26.6	37	24.8	63	28.0	31	17.1
	35 to 44	185	28.5	5	29.4	42	33.9	45	30.2	66	29.3	61	33.7
	45 to 54	124	19.1	5	29.4	25	20.2	38	25.5	58	25.8	41	22.7
	55 to 64	52	8.0	None	None	10	8.1	23	15.4	29	12.9	19	10.5
	65 or older	23	3.5	2	11.8	2	1.6	3	2.0	2	.9	15	8.3

그림 80. 빈 셀에 "없음"이 표시된 테이블

이제 테이블에서 네 개의 빈 셀에 값 0대신 없음이라는 텍스트가 표시됩니다.

결측 통계의 표시 값

통계를 계산할 수 없는 경우 기본값 표시 값은 시스템-결측값을 나타내는 데 사용되는 기호인 마침표(.)가 됩니다. 이 값은 "빈" 셀과는 다르므로 결측 통계의 표시 값과 케이스가 없는 셀의 표시 값은 별도로 제어됩니다.

1. 테이블 작성기를 엽니다(분석 메뉴, 표, 테이블 사용자 정의).
2. 변수 목록에서 하루 TV시청시간을 캔버스 분할창에서 열 영역 위쪽의 입금 수령 방법 위로 끌어 놓습니다.

하루 TV시청시간은 척도변수이므로 자동으로 통계 소스 변수가 되고 요약 통계가 평균으로 바뀝니다.

3. 캔버스 분할창의 테이블 미리보기에서 하루 TV시청시간을 마우스 오른쪽 단추로 클릭하고 팝업 컨텍스트 메뉴에서 요약 통계를 선택합니다.
4. 통계 목록에서 유효수를 선택하고 화살표 키를 눌러 출력 목록에 추가합니다.
5. 선택한 항목에 적용을 클릭합니다.
6. 옵션 탭을 클릭합니다.
7. 계산할 수 없는 통계의 텍스트 필드에 해당 없음을 입력합니다.
8. 확인을 클릭하여 테이블을 작성합니다.

		Hours per day watching TV											
		How get paid last week											
		Hourly wage		Daily wage		Weekly wage		Monthly salary		Annual salary		Other pay rate	
		Mean	Valid N	Mean	Valid N	Mean	Valid N	Mean	Valid N	Mean	Valid N	Mean	Valid N
Age category	Less than 25	3	71	NA	None	3	10	2	3	2	6	2	8
	25 to 34	3	134	5	2	2	30	2	29	2	52	2	22
	35 to 44	3	136	2	5	3	30	2	34	2	47	3	46
	45 to 54	2	90	2	4	2	22	2	36	2	45	2	34
	55 to 64	3	40	NA	None	3	7	2	15	2	23	3	15
	65 or older	3	18	2	2	1	1	NA	0	1	2	3	11

그림 81. 결측 통계에 대해 "해당 없음"이 표시된 테이블

테이블의 세 개의 셀 평균에 해당 없음이라는 텍스트가 표시됩니다. 각 케이스에서 해당 유효수 값을 보면 평균을 계산할 케이스가 없음을 알 수 있습니다.

하지만 케이스가 없는 셀에 표시되어야 하는 이 세 개의 유효수 값 중 하나가 없음이라는 레이블 대신 0으로 표시된다는 점이 약간 모순이 있어 보입니다. 이는 평균을 계산할 유효한 케이스는 없지만 이 범주가 실제로 비어있지는 않기 때문입니다. 두 개의 범주형 변수만 있는 원래 표로 되돌리면 이 교차 분석표 범주에 실제로 세 개의 케이스가 있음을 알 수 있습니다. 하지만 이 세 개의 케이스에 척도변수 하루 TV 시청 시간에 대한 결측값이 있으므로 유효한 케이스는 없습니다.

제 14 장 표본 파일

제품에 설치된 표본 파일은 설치 디렉토리 표본 추출 하위 디렉토리에서 찾으실 수 있습니다. 영어, 프랑스어, 독일어, 이탈리아어, 일본어, 한국어, 폴란드어, 러시아어, 중국어 간체, 스페인어 및 중국어 번체 등의 각 언어에 대해 Samples 하위 디렉터리 내에 별도의 폴더가 있습니다.

모든 언어에서 모든 표본 파일을 사용할 수 있는 것은 아닙니다. 특정 언어로 표본 파일을 사용할 수 없는 경우 해당 언어 폴더에 영어 버전의 표본 파일이 들어 있습니다.

설명

이것은 문서 전반에서 다양한 예제에 사용된 표본 파일에 대한 요약 설명입니다.

- **accidents.sav.** 이것은 주어진 지역에서 자동차 사고에 대한 연령 및 성별 위험 요인을 연구하는 보험회사와 관련이 있는 가설 데이터 파일입니다. 각 케이스는 연령 범주 및 성별의 교차 분류에 대응합니다.
- **adl.sav.** 이것은 뇌졸중 환자에 대해 제시된 치료 유형의 장점을 판단하려는 노력과 관련이 있는 가설 데이터 파일입니다. 의사는 여성 뇌졸중 환자를 두 그룹 중 하나에 무작위 할당합니다. 첫 번째로 받은 표준 물리치료, 두 번째로 받은 추가 정서적 치료. 치료 후 석달 동안 각 환자가 일상생활에서 흔한 활동들을 수행하는 기능을 순서 변수로써 스코어를 매겼습니다.
- **advert.sav.** 이것은 광고에 소비된 돈과 그 결과로써 판매에 소비된 돈 사이의 관계를 조사하려는 유통업체의 시도와 관련이 있는 가설 데이터 파일입니다. 이를 위해 과거 매출 수치와 관련 광고비를 수집했으며,
- **aflatoxin.sav.** 이것은 옥수수 작물에 대해서 농작물마다 농도가 크게 다른 독성 물질인 아플라톡신 검정과 관련이 있는 가설 데이터 파일입니다. 어느 곡물 가공업자가 8가지 농작물 중에서 16개 표본을 얻어 10억 분의 1단위로 아플라톡신 레벨을 측정하였습니다.
- **anorectic.sav** 연구원들은 식욕 감퇴/폭식증 등 표준적인 증상에 대한 연구와 병행하여 ¹ 알려진 섭식 장애를 갖고 있는 55명의 젊은이에 대한 연구를 수행하였습니다. 각 환자를 4년 동안 네 차례 진찰하여 총 220개의 관측값을 얻었습니다. 환자들을 대상으로 관찰하는 동안 16개 증상 각각에 대해 스코어를 매겼습니다. 환자 71번의 2회차, 환자 76번의 2회차, 환자 47번의 3회차 진료에 대한 증상 스코어가 누락되었으며 나머지 217개 관측값은 유효합니다.
- **bankloan.sav.** 이것은 은행이 기본 대부율을 낮추려는 시도와 관련이 있는 가설 데이터 파일입니다. 파일에는 과거 및 장래 전망 있는 850명의 고객에 대한 금융 정보 및 인구 통계학 정보가 들어 있습니다. 처음 700개의 케이스는 이전에 대출을 받은 고객입니다. 마지막 150 케이스는 은행이 신용 위험이 낮거나 높은 것으로 분류할 필요가 있는 장래 전망 있는 고객들입니다.
- **bankloan_binning.sav.** 이것은 과거 5,000명의 고객들에 관한 금융 정보 및 인구 통계학 정보가 들어있는 가설 데이터 파일입니다.

1. Van der Ham, T., J. J. Meulman, D. C. Van Strien, and H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, 363-368.

- **behavior.sav.** 고전적인 예제 ²에서 52명의 학생들에게 부탁하여 15가지 상황과 15가지 행동의 조합에 대해 0 = "매우 적절함"에서 9 = "매우 부적절함"까지 10단계 척도로 순위를 매기도록 요청했습니다. 각 응답자에 대해 평균을 구한 값은 상이도로 간주됩니다.
- **behavior_ini.sav.** 이 데이터 파일에는 *behavior.sav*에 대한 2차원적 솔루션의 초기 설정이 들어 있습니다.
- **brakes.sav.** 이것은 성능이 높은 자동차용 디스크 브레이크 생산 공장의 품질 관리와 관련이 있는 가설 데이터 파일입니다. 데이터 파일에는 생산 기계 8대 각각의 16개 디스크에 대한 지름 치수가 들어 있습니다. 브레이크의 목표 지름은 322mm입니다.
- **breakfast.sav.** 고전적인 연구³에서 21명의 Wharton School MBA 학생들과 그 배우자들에게 15가지 아침 식사 메뉴에 선호도에 따라 1부터 15까지 순위를 매기도록 부탁했습니다. 1은 '가장 좋아함'을 의미하고 15는 '가장 싫어함'을 의미합니다. "전반적인 선호"에서 "음료수만 포함된 가벼운 식사"에 이르는 서로 다른 여섯 가지 시나리오에 따라 응답자들의 선호도를 기록했습니다.
- **breakfast-overall.sav.** 이 데이터 파일에는 최초의 시나리오 "전반적인 선호도"에 대한 아침 식사 메뉴 중 좋아하는 음식이 들어 있습니다.
- **broadband_1.sav.** 이것은 지역별 국가 광대역 서비스 가입자 수가 들어 있는 가설 데이터 파일입니다. 데이터 파일에는 4년 동안 85개 지역에 대한 매달 가입자 수가 들어 있습니다.
- **broadband_2.sav.** 이 데이터 파일은 *broadband_1.sav*와 동일하나 추가 3달에 대한 데이터가 들어 있습니다.
- **car_insurance_claims.sav.** 자동차 손해배상 청구와 관련하여 ⁴ 외에서 제시 및 분석된 데이터 세트입니다. 평균 청구 금액은 평균 종속 변수를 보험 계약자의 연령, 차종 및 차량 연식과 관련시켜 설명하는 역 연결 함수를 사용하여 감마 분포로써 모형화할 수 있습니다. 기록에 남긴 청구 개수는 척도 가중치로 사용될 수 있습니다.
- **car_sales.sav.** 이 데이터 파일은 다양한 메이커와 차량 모형에 대한 예상 판매량, 정가 및 물리적 사양을 포함합니다. 정가와 물리적 사양은 *edmunds.com*과 제조업체 사이트에서 번갈아 입수하였습니다.
- **car_sales_uprepared.sav.** *car_sales.sav*의 수정된 버전으로서 변환된 버전의 필드가 포함되지 않습니다.
- **carpet.sav.** 자주 인용되는 예제인 ⁵에서는 새로운 카펫 세척제를 마케팅하려는 한 회사가 패키지 계획, 브랜드 이름, 값, 보증 유무 및 환불 유무라는 다섯 가지 요인이 고객 선호도에 어떤 영향을 미치는지 조사합니다. 패키지 계획에는 세 가지 요인 수준이 있고 그 각각은 세척용 솔의 위치에서 차이가 납니다. 그밖에도 세 가지 브랜드 이름(*K2R*, *Glory*, *Bissell*)이 있으며, 값 수준도 세 가지이고, 마지막 두 요인에 대해서는 각각 두 가지 수준(예 또는 아니오)이 있습니다. 이들 요인으로 정의된 22개의 프로파일에 대해 10명의 고객이 순위를 매겼습니다. 변수 선호도에는 각 프로파일의 평균 순위가 포함되어 있습니다. 낮은 순위는 선호도가 높음을 의미합니다. 이 변수는 각 프로파일에 대한 선호도의 전반적인 측도를 반영합니다.

2. Price, R. H., and D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579-586.

3. Green, P. E., and V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

4. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

5. Green, P. E., and Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

- **carpet_prefs.sav.** 이 데이터 파일은 *carpet.sav*에 설명된 것과 동일한 예제를 기반으로 하지만 10명의 고객 각각으로부터 수집한 실제 순위를 포함합니다. 고객들에게 가장 선호하는 제품부터 가장 덜 선호하는 제품에 이르기까지 22개 제품 프로파일에 순위를 매기도록 요청하였습니다. 변수 *PREF1*부터 *PREF22*까지는 *carpet_plan.sav*에 정의된 관련 프로파일의 식별자를 포함합니다.
- **catalog.sav.** 이 데이터 파일에는 카탈로그 회사에서 판매한 세 가지 제품의 가설적인 월별 판매 수치가 들어 있습니다. 가능한 5 가지 예측자 변수에 대한 데이터도 포함되어 있습니다.
- **catalog_seasfac.sav.** 이 데이터 파일은 부수적인 날짜 변수와 함께 계절 분해 프로시저에서 계산한 계절 요인 그룹을 추가한 것을 제외하고는 *catalog.sav*와 동일합니다.
- **cellular.sav.** 이것은 서비스 제공 사업자 이탈을 줄이려는 이동전화 회사의 시도와 관련이 있는 가설 데이터 파일입니다. 서비스 제공 사업자 이탈 성향 스코어가 0 ~ 100 범위로 계정에 적용됩니다. 스코어가 50 이상인 계정은 서비스 제공업체를 변경할 것을 기대할 수 있습니다.
- **ceramics.sav.** 이것은 새로운 프리미엄급 합금이 표준 합금보다 내열성이 훨씬 크지 여부를 판단하기 위한 제조업체의 시도와 관련이 있는 가설 데이터 파일입니다. 각 케이스는 합금 중 한 가지에 대한 별도의 검정을 나타냅니다. 베어링이 파손되는 온도를 기록합니다.
- **cereal.sav.** 이것은 880명의 사람들을 대상으로 연령, 성별, 결혼상태 및 활동적인 라이프스타일을 갖고 있는지 여부(1주에 적어도 두 번은 운동을 하고 있는지)와 함께 아침 식사 메뉴 선호도를 조사한 여론조사와 관련이 있는 가설 데이터 파일입니다. 각 케이스는 개별 반응자를 나타냅니다.
- **clothing_defects.sav.** 이것은 의류 공장에서 품질 관리 프로세스와 관련이 있는 가설 데이터 파일입니다. 공장에서 생산된 각 로트로부터 검사자들은 의복 표본을 선택하고 채택이 불가능한 의복 개수를 셉니다.
- **coffee.sav.** 이 데이터 파일은 여섯 가지 아이스커피 브랜드에 대해 인지된 이미지에 관한 데이터 파일입니다⁶. 23개의 아이스커피 이미지 속성 각각에 대해 해당 속성으로 설명할 수 있는 브랜드를 모두 선택하도록 사람들에게 요청했습니다. 브랜드 이름이 노출되지 않도록 여섯 개의 브랜드에는 AA, BB, CC, DD, EE 및 FF라고 이름을 붙였습니다.
- **contacts.sav.** 이것은 기업 컴퓨터 판매 대리점 그룹의 담당자 목록과 관련이 있는 가설 데이터 파일입니다. 각 담당자는 이들이 일하는 회사의 부서 및 그들의 회사내 지위별로 범주가 나누어졌습니다. 또한 마지막 판매량, 최종 판매 이후 시간 및 담당자 회사의 규모를 기록하였습니다.
- **creditpromo.sav.** 이것은 최근 신용카드 판촉의 효과를 평가하기 위한 백화점의 시도와 관련이 있는 가설 데이터 파일입니다. 이러한 목적을 위하여 500명의 카드 소지자를 무작위로 선택하였습니다. 이중 절반은 다음 세 달 동안 구매에 대한 이자율을 줄여 주겠다는 판촉 광고물을 받았습니다. 절반은 표준적인 시즌 광고물을 받았습니다.
- **customer_dbase.sav.** 이것은 데이터 웨어하우스의 정보를 사용하여 답변할 가능성이 가장 높은 고객들에게 특가 제공을 지원하려는 회사의 시도와 관련 있는 가설 데이터 파일입니다. 고객 층의 서브세트를 무작위로 선택하여 특가 제공을 지원하고 그들의 반응을 기록하였습니다.
- **customer_information.sav.** 이름 및 주소와 같은 고객 메일링 정보가 들어 있는 가설 데이터 파일입니다.
- **customer_subset.sav.** *customer_dbase.sav*의 80개 케이스의 서브세트입니다.

6. Kennedy, R., C. Riquier, and B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56-70.

- **debate.sav.** 이것은 정치적 토론의 참가자들을 대상으로 한 토론 전후 설문조사에서 대응반응과 관련 있는 가설 데이터 파일입니다. 각 케이스는 개별 반응자에 해당합니다.
- **debate_aggregate.sav.** 이것은 *debate.sav*의 반응을 통합하는 가설 데이터 파일입니다. 각 케이스는 토론 전후 선호도에 대한 교차 분류에 해당합니다.
- **demo.sav.** 이것은 매달 메일을 발송할 목적으로 구매 고객 데이터베이스와 관련 있는 가설 데이터 파일입니다. 다양한 인구 통계학 정보와 함께 고객이 제공에 반응했는지 여부가 기록됩니다.
- **demo_cs_1.sav.** 이것은 설문조사 정보 데이터베이스를 컴파일하려는 회사의 시도중 첫 단계와 관련 있는 가설 데이터 파일입니다. 각 케이스는 서로 다른 도시에 해당하며 지역, 지방, 구 및 도시 식별이 기록됩니다.
- **demo_cs_2.sav.** 이것은 설문조사 정보 데이터베이스를 컴파일하려는 회사의 시도중 두 번째 단계와 관련 있는 가설 데이터 파일입니다. 각 케이스는 첫 번째 단계에서 선정된 도시의 서로 다른 가구 단위에 해당하며 지역, 지방, 구, 도시, 구역 및 단위 식별이 기록됩니다. 계획 중 처음 두 단계의 표본 추출 정보도 포함됩니다.
- **demo_cs.sav.** 이것은 복잡한 표본추출 계획을 사용하여 수집한 설문조사 정보를 포함하고 있는 가설 데이터 파일입니다. 각 케이스는 서로 다른 가구 단위에 해당하며 다양한 인구 통계학적 정보와 표본 추출 정보가 기록됩니다.
- **diabetes_costs.sav.** 보험 회사가 당뇨병이 있는 보험 계약자에 대해 유지보수하는 정보가 포함된 가설 데이터 파일입니다. 각 케이스는 서로 다른 보험 가입자에 해당합니다.
- **dietstudy.sav.** 이것은 "스틸만 다이어트"의 연구 결과를 포함하는 가설 데이터 파일입니다.⁷ 각 케이스는 별도의 주제에 해당하며 파운드 단위로 다이어트 전후 체중을 기록하고 트리글리세이드 레벨을 mg/100 ml 단위로 기록합니다.
- **dmdata.sav.** 다이렉트 마케팅 회사의 인구 통계학 및 구매 정보가 들어 있는 가설 데이터 파일입니다. *dmdata2.sav*에는 검정 메일링을 수신한 연락처 서브세트에 대한 정보가 들어 있으며, *dmdata3.sav*에는 검정 메일을 수신하지 않은 나머지 연락처에 대한 정보가 들어 있습니다.
- **dvdplayer.sav.** 이것은 새 DVD 플레이어 개발과 관련 있는 가설 데이터 파일입니다. 마케팅 팀은 프로토타입을 사용하여 집중 그룹 데이터를 수집하였습니다. 각 케이스는 설문조사에 응한 각 사용자에게 해당하며 이들에 관한 몇 가지 인구 통계학적 정보와 프로토타입에 관한 앙케이트 반응을 기록합니다.
- **german_credit.sav.** 이 데이터 파일은 Irvine의 캘리포니아 대학⁸ Repository of Machine Learning 데이터베이스의 "독일인의 신용(German credit)" 데이터 세트에서 추출했습니다.
- **grocery_1month.sav.** 이것은 매주 "종합" 구매에 대한 *grocery_coupons.sav* 데이터 파일내의 가설 데이터 파일입니다. 따라서 각 케이스는 각 고객에 해당합니다. 결과적으로 매주 변하는 변수중 일부는 사라지며 이제 기록된 소비 금액은 4주의 연구기간중 소비한 금액의 합계가 됩니다.

7. Rickman, R., N. Mitchell, J. Dingman, and J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228., 54-58.

8. Blake, C. L., and C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

- **grocery_coupons.sav.** 이것은 고객의 구매 습관에 관심 있는 식료품 체인점에서 수집한 설문조사 데이터를 포함한 가설 데이터 파일입니다. 각 고객을 4주 동안 추적하며 각 케이스는 각각의 고객-주에 해당합니다. 고객이 그 주 동안 식료품 상점에서 소비한 금액을 포함하여 어디서 어떻게 쇼핑했는지에 관한 정보를 기록합니다.
- **guttman.sav.** Bell⁹은 가능한 사회 그룹을 보여 주는 테이블을 제안했습니다. Guttman¹⁰은 이 테이블의 일부를 사용했습니다. 그 중 사회적 상호작용, 그룹 소속감, 멤버의 친밀감, 관계의 형식성 등을 설명하는 다섯 개의 변수는 일곱 개의 이론적인 사회 그룹과 교차합니다. 이 일곱 개의 그룹에는 군중(예: 축구 시합 관중), 청중(예: 극장 관객 또는 강연회 방청객), 대중(예: 신문 구독자 또는 TV 시청자), 폭도(군중과 비슷하지만 상호작용이 훨씬 더 강한 그룹), 일차그룹(친밀한 관계), 이차그룹(자발적 관계), 현대 공동체(강한 친밀감과 특별한 서비스에 대한 필요의 결과로 발생한 느슨한 연합)가 있습니다.
- **health_funding.sav.** 이것은 의료보험 기금(모집단 100명당 금액), 발병율(모집단 10000명당 비율), 의료보험 제공업체 방문자(모집단 10000명당 비율)에 대한 데이터를 포함하는 가설 데이터 파일입니다. 각 케이스는 서로 다른 도시를 나타냅니다.
- **hivassay.sav.** 이것은 HIV 감염 여부를 발견하기 위한 신속한 검사 분석표를 개발하기 위한 제약 실험실의 시도와 관련 있는 가설 데이터 파일입니다. 검사 분석표 결과는 적색의 8단계 농도로 표시됩니다. 농도가 짙을수록 감염 가능성이 더 높음을 나타냅니다. 실험실에서 2000명의 혈액 표본에 대해 시험한 결과 이중 절반이 HIV에 감염되었으며 절반은 깨끗했습니다.
- **hourlywagedata.sav.** 이것은 다양한 경력을 갖춘 진료실 및 병원의 간호사의 시간당 급여에 관한 가설 데이터 파일입니다.
- **insurance_claims.sav.** 의심스럽거나, 사기성이 있는 청구를 거부하는 데 사용할 모형을 생성하려는 보험 회사를 위한 가설 데이터입니다. 각 케이스는 개별 청구를 나타냅니다.
- **insure.sav.** 이것은 고객이 10년만기 생명보험 계약에서 보험료를 청구할지 여부를 나타내는 위험 요인을 연구하는 보험 회사에 관한 가설 데이터 파일입니다. 데이터 파일의 각 케이스는 대응하는 계약을 나타내며, 그 중 한 계약은 연령과 성별에 따라 보험료 청구를 기록하였으며 다른 한 계약은 기록하지 않았습니다.
- **judges.sav.** 300개 체조 연기에 대해 훈련된 심판(+ 1명의 팬)이 매긴 스코어에 관한 가설 데이터 파일입니다. 각 행은 각각의 연기를 나타내며, 심판은 동일한 연기를 보았습니다.
- **kinship_dat.sav.** Rosenberg와 Kim¹¹은 친족 관계를 나타내는 15개의 단어(고모, 형제, 사촌, 딸, 아버지, 손녀, 할아버지, 할머니, 손자, 어머니, 조카, 조카딸, 누이, 아들, 삼촌)에 대한 조사를 착수했습니다. 조사자는 네 개의 대학생 그룹(여학생 두 그룹, 남학생 두 그룹)에 부탁하여 유사성을 기준으로 이들 단어를 정렬하도록 요청했습니다. 두 그룹(여학생 한 그룹, 남학생 한 그룹)에는 단어를 두 차례에 걸쳐 정렬하되 두 번째 정렬 기준은 첫 번째 경우와 다른 기준을 사용하도록 했습니다. 이런 식으로 6개 “소스”의 총계를 얻었

9. Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

10. Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469-506.

11. Rosenberg, S., and M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

습니다. 각 소스는 15 x 15 근접 행렬에 대응합니다. 이 행렬의 셀은 소스에 포함된 사람 수에서 오브젝트가 해당 소스에서 함께 분할된 횟수를 뺀 값과 같습니다.

- **kinship_ini.sav.** 이 데이터 파일에는 *kinship_dat.sav*에 대한 3차원적 솔루션의 초기 설정이 들어 있습니다.
- **kinship_var.sav.** 이 데이터 파일에는 *kinship_dat.sav*의 해법 차수를 해석하는 데 사용될 수 있는 독립변수 *gender*(성별), *gener*(세대) 및 *degree* (촌수)가 포함되어 있습니다. 특히, 이들 독립변수는 이러한 변수의 선형 조합으로 해법 공간을 제한하는 데 사용할 수 있습니다.
- **marketvalues.sav.** 이 데이터 파일은 1999 – 2000년 동안 Algonquin, Ill. 지역의 신규 주택 개발의 판매량에 관한 데이터 파일입니다. 이 판매량은 공식 레코드에 남아 있는 사항입니다.
- **nhis2000_subset.sav.** NHIS(National Health Interview Survey)는 미국 일반 시민을 대상으로 한 대규모의 인구 기반 설문조사입니다. 인터뷰는 국가 전역에서 뽑은 대표적인 표본 세대를 직접 대면하는 방식으로 진행합니다. 각 세대 멤버와의 인터뷰를 통해 각자의 생활 습관 및 건강 상태에 대한 관측값과 인구 통계학적 정보를 수집합니다. 이 데이터 파일은 2000년 설문조사의 정보 서브세트를 포함합니다. 미국 국립 건강 통계 센터 2000년 미국 국민 건강 설문조사, 공용 데이터 파일 및 문서자료. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/. 2003년 액세스됨.
- **ozone.sav.** 이 데이터에는 남은 변수로부터 오존 농도를 예측하기 위해 6가지 기상학적 변수에 관한 330개 관측값이 포함되어 있습니다. 이전의 조사¹², 무엇보다도¹³에서 이들 변수 간에 비선형성을 발견했으며 이는 표준 회귀분석 방법을 사용하는 데 걸림돌이 되고 있습니다.
- **pain_medication.sav.** 이것은 만성 관절염 통증 치료를 위한 항염증제 임상 시험 결과를 포함한 가설 데이터 파일입니다. 특히 관심을 받는 것은 약물이 효능을 발휘하는 데 소요되는 시간이며 기존의 약제와 비교하는 방법입니다.
- **patient_los.sav.** 이것은 심근경색(MI, 또는 심장마비)이 의심되어 병원에 내원한 환자들을 대상으로 한 치료 레코드가 포함된 가설 데이터 파일입니다. 각 케이스는 각 환자에 해당하며 환자의 입원에 관한 수많은 변수를 기록합니다.
- **patlos_sample.sav.** 이것은 심근경색(MI, 또는 심장마비)을 치료하는 동안 혈전 용해제를 투여받은 환자 표본에 대한 치료 레코드가 포함된 가설 데이터 파일입니다. 각 케이스는 각 환자에 해당하며 환자의 입원에 관한 수많은 변수를 기록합니다.
- **poll_cs.sav.** 이것은 입법 전 법안에 대한 공공 지지도를 판단하려는 여론 설문조사원의 시도에 관한 가설 데이터 파일입니다. 케이스는 등록된 유권자에 해당합니다. 각 케이스는 유권자가 살고 있는 군, 읍, 지구를 기록합니다.
- **poll_cs_sample.sav.** 이 가설 데이터 파일에는 *poll_cs.sav*에 등록된 유권자 표본이 들어 있습니다. 표본은 *poll.csplan* 계획 파일에 지정된 계획에 따라 선택되었으며 포함 확률 및 표본 가중치를 기록합니다. 그러나 표본추출 계획은 PPS(크기 비례) 방법을 사용하기 때문에 결합 선택 확률을 포함한 파일도 있음을 유의하

12. Breiman, L., and J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580-598.

13. Hastie, T., and R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

십시오(*poll_jointprob.sav*). 유권자 인구 통계학 정보에 해당하는 추가 변수 및 제안된 법안에 관한 그들의 의견을 수집하였으며 표본을 추출한 후 데이터 파일에 추가하였습니다.

- **property_assess.sav.** 이것은 제한된 자원에 대해 현재까지의 자산 가치를 평가하려는 카운티 평가자의 시도와 관련 있는 가설 데이터 파일입니다. 케이스는 작년 카운티에서 판매된 자산에 해당합니다. 데이터 파일의 각 케이스는 자산이 있는 군/구, 자산을 최종 방문한 평가자, 평가 후 경과 시간, 그당시 감정 결과, 자산에 대한 판매 가치를 기록합니다.
- **property_assess_cs.sav.** 이것은 제한된 자원에 대해 현재까지의 자산 가치를 평가하려는 주 평가자의 시도와 관련 있는 가설 데이터 파일입니다. 케이스는 주의 자산에 해당합니다. 데이터 파일의 각 케이스는 자산이 있는 군, 읍, 지구와 최종 평가 이후 경과된 시간, 그당시 감정 결과를 기록합니다.
- **property_assess_cs_sample.sav.** 이 가설 데이터 파일에는 *property_assess_cs.sav*에 등록된 자산 표본이 들어 있습니다. 표본은 *property_assess_csplan* 계획 파일에 지정된 계획에 따라 선택되었으며 포함 확률 및 표본 가중치를 기록합니다. 추가 변수 현재 값이 수집되었으며 표본을 선택한 후 데이터 파일에 추가되었습니다.
- **recidivism.sav.** 이것은 관할 영역에서 전과자 비율을 파악하려는 정부 법률 집행 기관의 시도와 관련 있는 가설 데이터 파일입니다. 각 케이스는 이전의 범죄자에 해당하며 그들의 인구 통계학 정보, 첫 번째 범죄 상세 내용, 첫 번째 구속 후 2년내에 범죄가 재발한 경우 두 번째 구속까지 경과된 시간을 기록합니다.
- **recidivism_cs_sample.sav.** 이것은 관할 영역에서 전과자 비율을 파악하려는 정부 법률 집행 기관의 시도와 관련 있는 가설 데이터 파일입니다. 각 케이스는 2003년 6월 동안 처음 구속된 후 출감한 이전의 범죄자에 해당하며 그들의 인구 통계학 정보, 첫 번째 범죄 상세 내용, 2006년 6월 말까지 범죄가 재발한 경우 두 번째 구속까지 경과된 시간을 기록합니다. 범죄자들은 *recidivism_cs_csplan*에 지정된 표본추출 계획에 따라 표본 추출된 부서에서 선택되었습니다. 왜냐하면 표본추출 계획은 PPS(크기 비례) 방법을 사용하므로 결합 선택 확률을 포함한 파일(*recidivism_cs_jointprob.sav*)도 있기 때문입니다.
- **rfm_transactions.sav.** 구매 데이터, 구매한 항목 및 각 트랜잭션 금액이 포함된 구매 트랜잭션 데이터가 있는 가설 데이터 파일입니다.
- **salesperformance.sav.** 이것은 두 가지 새 영업 교육 강의 평가와 관련 있는 가설 데이터 파일입니다. 60명의 직원들을 세 그룹으로 구분하였고 모두에게 표준 교육을 받게 하였습니다. 또한 그룹 2에는 기술 교육을, 그룹 3에는 실무 지습서를 제공합니다. 교육 강의를 끝나면 각 피고용자를 검정하여 스코어를 기록합니다. 데이터 파일의 각 케이스는 개별 훈련생을 나타내며 할당된 그룹과 시험 성적을 기록합니다.
- **satisf.sav.** 4군데 상점에 입점한 유통업체에서 실시간 고객 만족도 설문조사에 관한 가설 데이터 파일입니다. 582명의 고객 모두에 대해 설문조사를 실시하였으며 각 케이스는 한 명의 고객으로부터 받은 반응을 나타냅니다.
- **screws.sav.** 이 데이터 파일에는 나사, 볼트, 너트 및 못의 특성에 대한 정보가 포함되어 있습니다.¹⁴
- **shampoo_ph.sav.** 이것은 모발 관리 제품 공장의 품질 관리 프로세스와 관련이 있는 가설 데이터 파일입니다. 정기적인 시간 간격으로 6벌의 서로 다른 제품을 측정하여 pH를 기록하였습니다. 대상 범위는 4.5 – 5.5입니다.

14. Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.

- **ships.sav.** 파도에 의한 화물선박 손상에 관해 ¹⁵ 이외에서 제시 및 분석된 데이터 세트입니다. 사고 개수는 선박의 종류, 건조 기간, 서비스 기간을 가정한 포이송 비율에서 발생하는 것으로 모형화할 수 있습니다. 요인의 교차 분류에 의해 형성된 각 포 셀의 총 서비스 개월수는 위험에 대한 노출 값을 제공합니다.
- **site.sav.** 이것은 사업 확장을 위해 새 부지를 선택하려는 회사의 시도와 관련 있는 가설 데이터 파일입니다. 이 회사는 두 명의 컨설턴트를 고용하여 부지에 대한 전망을 따로따로 평가하게 하였고 확장 보고서에 서는 각 부지를 "우수", "양호" 또는 "불량"으로 요약하였습니다.
- **smokers.sav.** 이 데이터 파일은 1998년 전국 마약 남용에 관한 가정 설문조사에서 추출한 것으로 미국 가정의 확률 표본입니다. (<http://dx.doi.org/10.3886/ICPSR02934>) 따라서 이 데이터 파일 분석의 첫 단계는 인구 동향을 반영하는 데이터에 가중치를 더하는 작업이 되어야 합니다.
- **stocks.sav** 이 가설 데이터 파일에는 일 년 동안의 주가와 거래량이 들어 있습니다.
- **stroke_clean.sav.** 이것은 데이터 준비 옵션의 절차를 사용하여 삭제한 의료 데이터베이스의 상태를 포함한 가설 데이터 파일입니다.
- **stroke_invalid.sav.** 이 가설 데이터 파일은 의료 데이터베이스의 초기 상태와 여러 데이터 입력 오류를 포함합니다.
- **stroke_survival.** 이 가설 데이터 파일은 허혈성 뇌졸중 발생 후 수많은 도전과제에 직면하고 있는 재활 프로그램을 종료한 환자들의 생존 시간과 관련 있습니다. 뇌졸중, 심장마비의 발생, 허혈성 뇌졸중 또는 출혈성 뇌졸중으로 구분되며 이벤트 발생 시간이 기록됩니다. 표본은 현재까지 살아남은 사람들에 대한 데이터입니다. 왜냐하면 뇌졸중 후 시행한 재활 프로그램 마지막에 생존한 환자만 포함하기 때문입니다.
- **stroke_valid.sav.** 이것은 데이터 검증 절차를 사용하여 값을 점검한 후 의료 데이터베이스의 상태를 포함한 가설 데이터 파일입니다. 여기에는 여전히 잠재적인 이상항목 케이스가 포함됩니다.
- **survey_sample.sav.** 이 데이터 파일은 인구 통계학 데이터 및 다양한 태도 측도 등 설문조사 데이터가 포함됩니다. 이 데이터 파일은 1998 NORC General Social Survey의 변수 서브세트를 기준으로 하지만, 설명을 목적으로 일부 데이터 값이 수정되었으며 추가 가상 변수가 추가되었습니다.
- **tcm_kpi.sav.** 이것은 비즈니스의 매주 핵심성과지표(KPI) 값이 포함된 가설 데이터 파일입니다. 여기에는 또한 동일한 기간 동안 제어 가능한 다수의 메트릭에 대한 매주 데이터가 포함됩니다.
- **tcm_kpi_upd.sav.** 이 데이터 파일은 *tcm_kpi.sav*와 동일하나 추가 4주에 대한 데이터가 들어 있습니다.
- **telco.sav** 이것은 고객층에서 서비스 제공 사업자 이탈을 줄이려는 통신 회사의 시도와 관련이 있는 가설 데이터 파일입니다. 각 케이스는 개별 고객에 해당하며 다양한 인구 통계학 정보와 서비스 사용 정보가 기록됩니다.
- **telco_extra.sav.** 이 데이터 파일은 *telco.sav* 데이터 파일과 유사하지만 "tenure"(보유 기간) 및 로그 변환된 고객 소비 변수는 제거되었고 표준화된 로그 변환된 고객 소비 변수로 대체되었습니다.
- **telco_missing.sav.** 이 데이터 파일은 *telco.sav* 데이터 파일 서브세트지만 인구 통계학적 정보 데이터 값 중 일부는 결측값으로 대체되었습니다.
- **testmarket.sav.** 이 가설 데이터 파일은 메뉴에 새로운 품목을 추가하려는 패스트푸드 체인점의 계획과 관련이 있습니다. 신제품을 판촉하는 세 가지 가능한 캠페인이 있습니다. 따라서 무작위로 선택된 여러 시장

15. McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

위치에 새 품목이 도입됩니다. 각 위치에 서로 다른 판촉을 사용하며 처음 4주 동안 새 품목의 주간 판매량을 기록합니다. 각 케이스는 별도의 위치-주에 해당합니다.

- **testmarket_1month.sav.** 이것은 매주 "종합" 판매량에 대한 *testmarket.sav* 데이터 파일내의 가설 데이터 파일입니다. 따라서 각 케이스는 각 위치에 해당합니다. 결과적으로 매주 변하는 변수중 일부는 사라지며 이제 기록된 판매량은 4주의 연구기간중 판매량의 합계가 됩니다.
- **tree_car.sav.** 이것은 인구 통계학 및 차량 구매 가격 데이터를 포함한 가설 데이터 파일입니다.
- **tree_credit.sav.** 이것은 인구 통계학 및 은행 대부 이력을 포함한 가설 데이터 파일입니다.
- **tree_missing_data.sav** 이것은 수많은 결측값을 갖고 있는 인구 통계학 데이터 및 은행 대부 이력 데이터를 포함한 가설 데이터 파일입니다.
- **tree_score_car.sav.** 이것은 인구 통계학 및 차량 구매 가격 데이터를 포함한 가설 데이터 파일입니다.
- **tree_textdata.sav.** 측정 수준 및 값 레이블을 할당하기 전 주로 기본값 변수 상태를 보여주기 위한 목적으로 두 변수만을 포함한 단순 데이터 파일입니다.
- **tv-survey.sav.** 이것은 성공적인 프로그램 운영을 확장할지 여부를 고려하기 위해 텔레비전 스튜디오에서 실시한 여론설문조사에 관한 가설 데이터 파일입니다. 906명의 반응자에게 다양한 조건 하에서 프로그램을 시청할지 여부에 대해 질문하였습니다. 각 행은 개별 반응자를 나타내며 각 열은 개별 조건을 나타냅니다.
- **ulcer_recurrence.sav.** 이 파일에는 궤양의 재발 방지를 위한 두 가지 치료법의 효능을 비교하도록 계획된 연구의 정보 일부가 들어 있습니다. 이것은 구간 중도 절단 자료(interval-censored data)의 좋은 예를 제공하며 ¹⁶ 이외의 장소에서 제시 및 분석되었습니다.
- **ulcer_recurrence_recoded.sav.** 이 파일은 *ulcer_recurrence.sav*의 정보를 재정리하므로 단순히 연구가 끝난 후의 이벤트 확률이 아닌 연구의 각 구간에 대한 이벤트 확률을 모형화할 수 있습니다. 이것은 ¹⁷ 이외의 장소에서 제시 및 분석되었습니다.
- **verd1985.sav.** 이 데이터 파일은 설문조사 ¹⁸와 관련 있습니다. 15명의 사람들이 8개의 변수에 대해 반응한 내용을 기록했습니다. 관심 변수를 세 가지 세트로 구분하였습니다. 변수 세트 1에는 연령과 결혼상태가 포함되고, 변수 세트 2에는 애완동물과 신문이 포함되고, 변수 세트 3에는 음악과 거주가 포함됩니다. 애완동물은 다중 명목으로 척도화되고 연령은 순서로 척도화됩니다. 다른 모든 변수는 단일 명목으로 척도화됩니다.
- **virus.sav.** 이것은 바이러스가 네트워크에 미치는 영향을 판단하려는 인터넷 서비스 공급업체(ISP)의 시도와 관련 있는 가설 데이터 파일입니다. 발견 시점에서 위협이 제거된 시점까지의 시간 동안 자사의 네트워크에서 감염된 이메일 트래픽의 대략적인 퍼센트를 추적했습니다.
- **wheeze_steubenville.sav.** 이것은 대기 오염이 자녀들의 건강에 미치는 영향에 대한 종단적 연구의 서브세트입니다.¹⁹ 데이터에는 오하이오주 Steubenville의 7~10세 어린이들을 대상으로 연구 첫째 동안은 어머니가 흡연자인지 여부에 관한 레코드를 포함, 호흡기 천명에 대한 이원적 반복 측도 결과가 들어 있습니다.

16. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

17. Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

18. Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

19. Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, and B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, 366-374.

- **workprog.sav.** 이것은 불리한 조건을 가진 사람들을 더 나은 직업으로 배치하려는 정부 연구 프로그램에 관한 가설 데이터 파일입니다. 잠재적인 프로그램 참가자 표본을 추적하였습니다. 이들 중 일부는 프로그램 영입을 위하여 무작위로 선택된 반면 다른 이들은 그렇지 않습니다. 각 케이스는 개별 프로그램 참가자를 나타냅니다.
- **worldsales.sav** 이 가설 데이터 파일에는 대륙 및 제품별 판매 수입이 들어 있습니다.

주의사항

이 정보는 미국에서 제공되는 제품 및 서비스용으로 작성된 것입니다. 본 자료는 다른 언어로도 제공될 수 있습니다. 그러나 자료에 접근하기 위해서는 해당 언어로 된 제품 또는 제품 버전의 사본이 필요할 수 있습니다.

IBM은 다른 국가에서 이 책에 기술된 제품, 서비스 또는 기능을 제공하지 않을 수도 있습니다. 현재 사용할 수 있는 제품 및 서비스에 대한 정보는 한국 IBM 담당자에게 문의하십시오. 이 책에서 IBM 제품, 프로그램 또는 서비스를 언급했다고 해서 해당 IBM 제품, 프로그램 또는 서비스만을 사용할 수 있다는 것을 의미하지는 않습니다. IBM의 지적 재산권을 침해하지 않는 한, 기능상으로 동등한 제품, 프로그램 또는 서비스를 대신 사용할 수도 있습니다. 그러나 비IBM 제품, 프로그램 또는 서비스의 운영에 대한 평가 및 검증은 사용자의 책임입니다.

IBM은 이 책에서 다루고 있는 특정 내용에 대해 특허를 보유하고 있거나 현재 특허 출원 중일 수 있습니다. 이 책을 제공한다고 해서 특허에 대한 라이선스까지 부여하는 것은 아닙니다. 라이선스에 대한 의문사항은 다음으로 문의하십시오.

150-945

서울특별시 영등포구 국제금융로 10, 31FC

한국 아이.비.엠 주식회사

대표전화서비스: 02-3781-7114

2바이트(DBCS) 정보에 관한 라이선스 문의는 한국 IBM에 문의하거나 다음 주소로 서면 문의하시기 바랍니다.

Intellectual Property Licensing

Legal and Intellectual Property Law

IBM Japan Ltd.

19-21, Nihonbashi-Hakozakicho, Chuo-ku

Tokyo 103-8510, Japan

IBM은 타인의 권리 비침해, 상품성 및 특정 목적에의 적합성에 대한 묵시적 보증을 포함하여(단, 이에 한하지 않음) 묵시적이든 명시적이든 어떠한 종류의 보증 없이 이 책을 "현상태대로" 제공합니다. 일부 지역에서는 특정 상거래에 있어 명시적 또는 묵시적 보증 책임에 대한 불인정을 허용하지 않으므로, 이런 지역에서는 위의 사항이 적용되지 않습니다.

이 정보에는 기술적으로 부정확한 내용이나 인쇄상의 오류가 있을 수 있습니다. 이 정보는 주기적으로 변경되며, 변경된 사항은 최신판에 통합됩니다. IBM은 이 책에서 설명한 제품 및/또는 프로그램을 사전 통지 없이 언제든지 개선 및/또는 변경할 수 있습니다.

이 정보에서 언급되는 비IBM의 웹 사이트는 단지 편의상 제공된 것으로, 어떤 방식으로든 이들 웹 사이트를 옹호하고자 하는 것은 아닙니다. 해당 웹 사이트의 자료는 본 IBM 제품 자료의 일부가 아니므로 해당 웹 사이트 사용으로 인한 위험은 사용자 본인이 감수해야 합니다.

IBM은 귀하의 권리를 침해하지 않는 범위 내에서 적절하다고 생각하는 방식으로 귀하가 제공한 정보를 사용하거나 배포할 수 있습니다.

(i) 독립적으로 작성된 프로그램과 기타 프로그램(본 프로그램 포함) 간의 정보 교환 및 (ii) 교환된 정보의 상호 이용을 목적으로 본 프로그램에 관한 정보를 얻고자 하는 라이선스 사용자는 다음 주소로 문의하십시오.

150-945

서울특별시 영등포구 국제금융로 10, 3IFC

한국 아이.비.엠 주식회사

대표전화서비스: 02-3781-7114

이러한 정보는 해당 조건(예를 들면, 사용료 지불 등)하에서 사용될 수 있습니다.

이 정보에 기술된 라이선스가 부여된 프로그램 및 프로그램에 대해 사용 가능한 모든 라이선스가 부여된 자료는 IBM이 IBM 기본 계약, IBM 프로그램 라이선스 계약(IPLA) 또는 이와 동등한 계약에 따라 제공한 것입니다.

인용된 성능 데이터와 고객 예제는 예시 용도로만 제공됩니다. 실제 성능 결과는 특정 구성과 운영 조건에 따라 다를 수 있습니다.

비IBM 제품에 관한 정보는 해당 제품의 공급업체, 공개자료 또는 기타 범용 소스로부터 얻은 것입니다. IBM에서는 이러한 제품들을 테스트하지 않았으므로, 비IBM 제품과 관련된 성능의 정확성, 호환성 또는 기타 청구에 대해서는 확신할 수 없습니다. 비IBM 제품의 성능에 대한 의문사항은 해당 제품의 공급업체에 문의하십시오.

IBM이 제시하는 방향 또는 의도에 관한 모든 언급은 특별한 통지 없이 변경될 수 있습니다.

이 정보에는 일상의 비즈니스 운영에서 사용되는 자료 및 보고서에 대한 예제가 들어 있습니다. 이들 예제는 개념을 가능한 완벽하게 설명하기 위하여 개인, 회사, 상표 및 제품의 이름이 사용될 수 있습니다. 이들 이름은 모두 가공의 것이며 실제 기업 및 인물과 유사하더라도 이는 전적으로 우연입니다.

저작권 라이선스:

이 정보에는 여러 운영 플랫폼에서의 프로그래밍 기법을 보여주는 원어로 된 샘플 응용프로그램이 들어 있습니다. 귀하는 이러한 샘플 프로그램의 작성 기준이 된 운영 플랫폼의 애플리케이션 프로그래밍 인터페이스(API)에 부합하는 애플리케이션을 개발, 사용, 판매 또는 배포할 목적으로 IBM에 추가 비용을 지불하지 않고 이들 샘플 프로그램을 어떠한 형태로든 복사, 수정 및 배포할 수 있습니다. 이러한 샘플 프로그램은 모든 조건하에서 완전히 테스트된 것은 아닙니다. 따라서 IBM은 이러한 프로그램의 신뢰성, 서비스 가능성 또는 기능을 보증하거나 진술하지 않습니다. 본 샘플 프로그램은 일체의 보증 없이 "현상태대로" 제공됩니다. IBM은 귀하의 샘플 프로그램 사용과 관련된 손해에 대해 책임을 지지 않습니다.

이러한 샘플 프로그램 또는 파생 제품의 각 사본이나 그 일부에는 반드시 다음과 같은 저작권 표시가 포함되어야 합니다.

© (귀하의 회사명) (연도). 이 코드의 일부는 IBM Corp.의 샘플 프로그램에서 파생됩니다.

© Copyright IBM Corp. _enter the year or years_. All rights reserved.

상표

IBM, IBM 로고 및 ibm.com은 전세계 여러 국가에 등록된 International Business Machines Corp.의 상표 또는 등록상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 타사의 상표입니다. 현재 IBM 상표 목록은 "저작권 및 상표 정보" 웹 페이지(www.ibm.com/legal/copytrade.shtml)에 있습니다.

Adobe, Adobe 로고, PostScript 및 PostScript 로고는 미국 및/또는 기타 국가에서 사용되는 Adobe Systems Incorporated의 등록상표 또는 상표입니다.

Intel, Intel 로고, Intel Inside, Intel Inside 로고, Intel Centrino, Intel Centrino 로고, Celeron, Intel Xeon, Intel SpeedStep, Itanium 및 Pentium은 미국 또는 기타 국가에서 사용되는 Intel Corporation 또는 그 계열사의 상표 또는 등록상표입니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

Microsoft, Windows, Windows NT 및 Windows 로고는 미국 또는 기타 국가에서 사용되는 Microsoft Corporation의 상표입니다.

UNIX는 미국 및 기타 국가에서 사용되는 The Open Group의 등록상표입니다.

Java 및 모든 Java 기반 상표와 로고는 Oracle 및/또는 그 계열사의 상표 또는 등록상표입니다.

색인

[가]

값
범주 레이블 및 값 표시 59
값 및 값 레이블 59
개수
유효수 비교 62
검정 통계량
테이블 사용자 정의 19, 71
결측값 62, 93
테이블 사용자 정의에 포함 94
퍼센트 계산에 대한 영향 94
계산된 범주
소계에서 47
테이블 사용자 정의 15, 45
표시 형식 16
표현식에서 범주 감추기 46

교차 분석표
테이블 사용자 정의 23
그룹 요약
척도변수 65
그룹 총계 39

[나]

날짜
테이블 사용자 정의에 현재 날짜 포함 18
누적변수
누적 레이어 변수 35
다중 요약 통계 소스 변수 57
변수마다 다른 요약 통계 64
척도변수 61
테이블 사용자 정의 29

[다]

다중 응답 세트 85
다중 범주 세트의 중복 반응수 17
유의성 검증 83, 90, 91
퍼센트 9

[라]

레이블
요약 통계에 대한 레이블 텍스트 바꾸기 99
레이어 변수
누적 레이어 변수 35
레이어 테이블 인쇄 36
중첩 레이어 변수 35
테이블 사용자 정의 34, 35
레이어가 있는 테이블 인쇄 36

[마]

모드
테이블 사용자 정의 9

[바]

범위
테이블 사용자 정의 9
범주 값 표시 59
범주 다시 정렬
테이블 사용자 정의 12
범주 삭제
테이블 사용자 정의 12, 26
범주 생략
테이블 사용자 정의 26
범주 정렬
테이블 사용자 정의 26
범주 제외
테이블 사용자 정의 12, 26
범주 합치기
테이블 사용자 정의 42
변수 레이블
테이블 사용자 정의에 출력 없음 6
변수마다 다른 요약 통계
누적표 64
분산
테이블 사용자 정의 9
분할 파일 처리
테이블 사용자 정의 4
빈 셀
테이블 사용자 정의에 표시된 값 17, 100

빈도표

테이블 사용자 정의 16, 51

[사]

사용자 결측값 93
사용자 정의 총계 요약 통계 58
사후 계산된 범주
테이블 사용자 정의 15, 45
소계 41
소계 범주 감추기 42
테이블 사용자 정의 12, 37
소수점 이하 자릿수
테이블 사용자 정의에 표시된 소수점 이하 자릿수 제어 6, 25, 97
시간
테이블 사용자 정의에 현재 시간 포함 18
시스템-결측값 93
신뢰구간 69

[아]

열 너비
테이블 사용자 정의의 통제변수 17, 100
열 비율 통계
테이블 사용자 정의 77
열 평균 통계
테이블 사용자 정의 74
요약 통계 55
누적표 57
레이블 텍스트 바꾸기 99
변수마다 다른 요약이 표시된 누적표 64
사용자 정의 총계 요약 통계 58
소스 변수 56
소스 차원 56
표시 형식 97
요약 통계 소스 변수
척도변수 66
유의성 검증
다중 응답 세트 90, 91
테이블 사용자 정의 19
유효수 62, 94
테이블 사용자 정의 9

[자]

- 제목
 - 테이블 사용자 정의 18
- 중위수 62
 - 테이블 사용자 정의 9
- 중첩 변수
 - 척도변수 66
 - 테이블 사용자 정의 30, 32

[차]

- 척도변수
 - 그룹 요약 65
 - 누적 61
 - 다중 요약 통계 62
 - 요약 통계 61
 - 중첩 66
 - 행 및 열 범주형 변수에 따라 그룹화한 요약 65
- 총계
 - 그룹 총계 39
 - 레이어 40
 - 범주 제외 38
 - 중첩 테이블 39
 - 출력 위치 38
 - 테이블 사용자 정의 12, 23, 37
 - 테이블 사용자 정의에 대한 주변 합계 25
- 최대값
 - 테이블 사용자 정의 9
- 최소값
 - 테이블 사용자 정의 9
- 출력된 소수점 이하 자릿수 제어 25
- 측정 수준
 - 테이블 사용자 정의 바꾸기 1

[카]

- 카이제곱
 - 테이블 사용자 정의 71
- 캡션
 - 테이블 사용자 정의 18
- 코너 레이블
 - 테이블 사용자 정의 18

[타]

- 테이블
 - 테이블 사용자 정의 1
- 테이블 사용자 정의
 - 검정 통계량 19, 71
 - 계산된 범주 12, 15, 45
 - 공유 범주가 있는 변수에 대한 테이블 16, 51
 - 교차 분석표 23
 - 누적변수 29
 - 다중 범주 세트 17
 - 다중 응답 세트 1, 85
 - 다중 응답 세트의 퍼센트 9
 - 레이어 변수 34, 35
 - 레이어 테이블 인쇄 36
 - 범주 다시 정렬 12
 - 범주 정렬 26
 - 범주 제외 12, 26
 - 범주 합치기 42
 - 범주형 변수 1
 - 범주형 변수에 대한 값 레이블 1
 - 범주형 변수에 대한 단순 통계표 22
 - 변수 이름과 레이블 보이기 및 감추기 6
 - 분할 파일 처리 4
 - 빈 셀 17
 - 빈도표 16, 51
 - 사용자 정의 총계 10
 - 사후 계산된 범주 15, 45
 - 소계 12, 37
 - 소계 범주 감추기 42
 - 열 너비 17
 - 요약 통계 7, 8, 9
 - 요약 통계 차원 바꾸기 10
 - 요약 통계 표시 형식 11
 - 요약 통계에 대한 레이블 바꾸기 22
 - 유의성 검증과 다중 응답 83
 - 제목 18
 - 제외된 범주가 있는 테이블의 총계 26
 - 주변 합계 25
 - 중첩 레이어 변수 35
 - 중첩 변수 30, 32
 - 척도 요약에 대해 결측값 제외 17
 - 척도변수 1
 - 총계 12, 23, 37
 - 최소 보기 32
 - 출력된 소수점 이하 자릿수 제어 6
 - 측정 수준 변경 1

테이블 사용자 정의 (계속)

- 캡션 18
- 코너 레이블 18
- 테이블 작성 방법 4
- 통계 레이블 숨김 22
- 통계 소스 차원 24
- 퍼센트 7, 8, 22, 24
- 평균-빈도표 10
- 표시 형식 6
- 행 및 열 퍼센트 22
- 행과 열 변수 바꾸기 33
- comperimeter 테이블 16, 51
- 테이블 사용자 정의의 통계 레이블 감추기 22
- 통계
 - 누적표 57
 - 사용자 정의 총계 요약 통계 58
 - 요약 통계 55

[파]

- 퍼센트
 - 결측값 94
 - 다중 응답 세트 9
 - 테이블 사용자 정의 7, 8, 22, 24
- 평균 61
 - 테이블 사용자 정의 9
- 평균-빈도표 10, 58
- 표본 파일
 - 위치 103
- 표시 형식 25
 - 테이블 사용자 정의의 요약 통계 11, 97
- 표준 편차
 - 테이블 사용자 정의 9

[하]

- 하위 그룹 총계 39
- 합계
 - 테이블 사용자 정의 9
- 합계 N 94

C

- comperimeter 테이블 16, 51

