4-3. 데이터 핸들링 - I

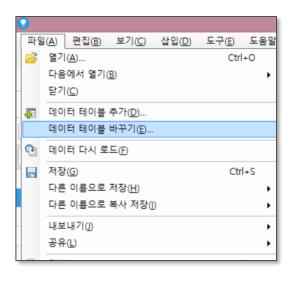
데이터 바꾸기/Join/행삽입/열삽입

목 차

- 1. 데이터 기본
 - 데이터 바꾸기 / 추가 / 다시로드
 - 데이터 테이블 속성
 - 컬럼 속성
 - 문서 속성
 - 마킹 데이터 핸들링
 - 삽입 > 계산된 컬럼 / 저장함에 넣은 컬럼 / 계층
- 2. 데이터 연결(Join)
- 3. 행 삽입 (Add Rows) : 삽입 > 행
- 4. 열 추가(Add Columns) : 삽입 > 컬럼

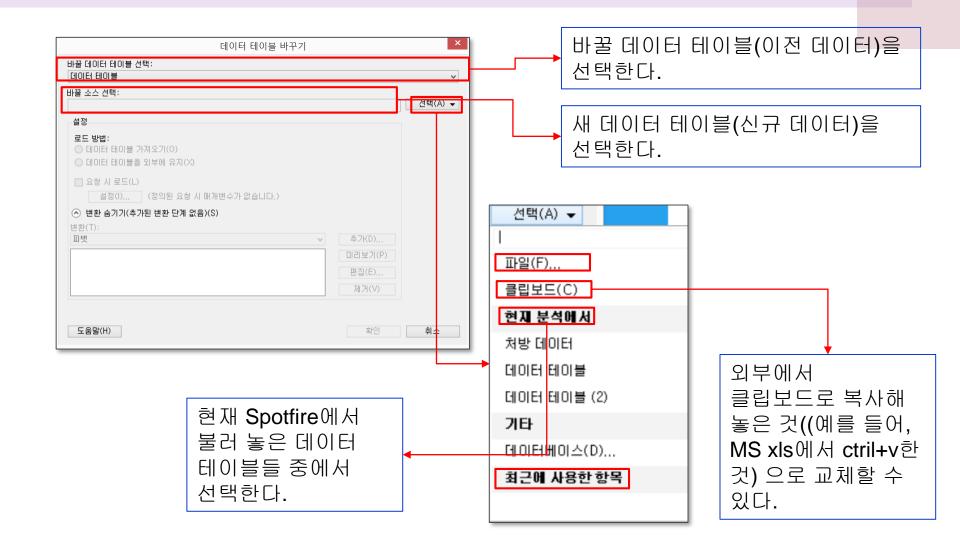
1. 데이터 기본 : 데이터 테이블 바꾸기

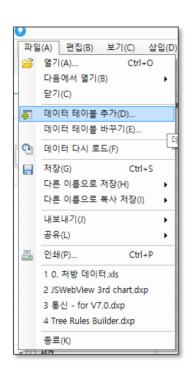
- Spotfire에서는 새 데이터가 이전 데이터와 형식(format)이 유사한 경우 문서의 시각화, 계산 및 설정을 재사용할 수 있다. 예를 들어, 매월 동일한 양식의 데이터를 이용하여 반복적인 업무를 하는 경우 이 기능은 매우 유용하다.
- 또는 지난 달에 판매 수치 등에 대한 분석을 만든 후에, 이어 이번 달의 분석을 다시 만들 때 유용하다.



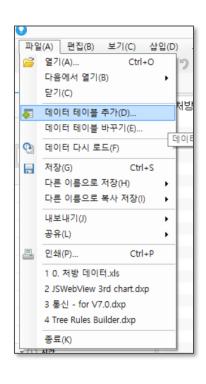
 먼저 1월부터의 데이터를 사용하여 전체 분석을 만들고, 시각화나 계산 등을 설정하고, 파일을 저장한다. 2월의 판매 수치를 이용할 수 있을 때 같은 파일을 다시 열어 1월의 데이터를 2월의 데이터로 바꾸면 1월의 시각화로 작성해 놓은 시각화가 2월의 데이터로 업데이트된다. 하지만 이를 위해서는 2월의 데이터 테이블이 1월의 데이터 테이블과 동일한 컬럼 이름 및 형식으로 되어 있어야만 한다.

1. 데이터 기본 : 데이터 테이블 바꾸기





- Spotfire에서는 여러 개의 데이터 테이블을 하나의 분석 파일(.dxp)에 가져와서 다양한 방법으로 여러 데이터를 분석에 사용할 수 있다.
- 이 메뉴는 처음 데이터를 가져 온 다음 두번째부터 추가로 데이터를 가져 오거나, 기존에 가져온 데이터의 형식을 변경하여 (예, Pivot, Unpivot) 기존 테이블 외에 추가로 테이블을 생성, 저장하여야 할 때 사용한다.
- 새로 데이터를 가져오면 시각화 좌측 상단에 좁은 색상의 막대(줄무늬; stripe)가 자동으로 생성되며, 이 부분은 여러 데이터 테이블을 구분하고, 필터를 사용할 때 혼동되지 않도록 하는데 사용한다.
- 만일 나중에 줄무늬 2개의 테이블을 연결(join)하면 두개의 테이블은 동일한 줄무늬를 표시하게 된다.



- Spotfire에서는 새 컬럼, 새 행, 새 데이터 테이블 등 다양한 방법으로 데이터를 분석에 추가할 수 있다.
- 새 데이터가 이전에 열린 데이터 테이블과 관련이 없거나 새 데이터의 형식이 다른 경우(피벗 및 언피벗) 데이터를 별도의 데이터 테이블로 추가하는 것이 좋다.
- 특정 데이터 테이블에서 시각화를 만들고 해당 데이터 테이블에 다른 데이터 테이블에서 만든 시각화에 적용하려는 필터링 및 마킹이 있는 경우 두 테이블 간의 관계를 정의해야 한다.
- 관계가 유용하려면 두 데이터 테이블 모두에서 사용 가능한 키 컬럼(지시자 컬럼)이 하나 이상 있고 해당 키 컬럼을 사용하여 두 번째 데이터 테이블의 행에 해당하는 첫 번째 데이터 테이블의 행을 정의해야 한다.
- 고유한 지시자를 설정하는 데 키 컬럼이 여러 개 필요한 경우 지시자 컬럼별로 하나의 관계를 추가해야 한다.

 이전 데이터와 새로 추가한 데이터간에 어떤 동일한 컬럼이 존재한다면 이 것을 키 컬럼(지시자 컬럼) 혹은 Join Key라고 하며, 그 2개의 데이터를 하나로 합칠 수 있다.

ID	이름	성별	나이	지역	최초 회원 가입일
pnd022	김도빈	남	38	서울	2008-02-10
pnd101	박진수	남	45	경기	2006-11-16
pnd278	최윤희	여	24	강원	2006-12-15
pnd171	이기영	남	33	전남	2007-01-04
pnd050	김지수	여	28	전북	2006-11-18
pnd196	이원균	남	40	경남	2006-12-20
pnd018	권진경	여	36	경북	2007-01-05
pnd148	유인나	여	50	충남	2007-01-02

'고객 정보' 데이터

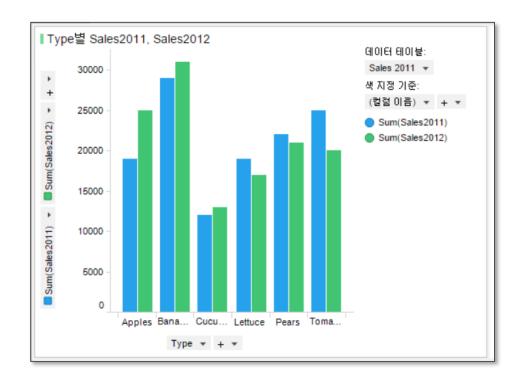
ID	이름	구매일	구매품목	모델명	가격	구입처	결제방법
pnd022	김도빈	2007-01-01	냉장고	SRT756BSLE	399000	백화점	현금
pnd101	박진수	2007-01-02	TV	UN46C8000XF	4450000	하이마트	카드-일시불
pnd278	최윤희	2007-01-03	휴대폰	SHW-M100S	970000	온라인	카드-일시불
pnd171	이기영	2007-01-04	세탁기	WR-PD179UB	1450000	백화점	카드-할부
pnd050	김지수	2007-01-05	냉장고	SRT746ABZF	3015000	백화점	현금/카드
pnd196	이원균	2007-01-06	세탁기	WR-HF137UD	1799000	하이마트	현금/카드
pnd018	권진경	2007-01-07	냉장고	SRL358US	770000	온라인	카드-할부
pnd148	유인나	2007-01-08	냉장고	SR482YC	1080000	백화점	카드-할부
						HAIRIE	

'판매 정보' 데이터

- 위의 예제 데이터에서는각각 'ID'와 '이름' 컬럼이 두 데이터 테이블에서 공통적으로 동일하게 사용되는 키 컬럼(지시자 컬럼)이다.
- 한 시각화에서 여러 데이터 테이블의 데이터를 결합하려면 데이터 테이블을 기존 대로 로드하지만 데이터 테이블의 데이터를 결합하기 위해 데이터 테이블 간에 하나 이상의 컬럼이 일치해야 한다. **컬럼이 일치하려면 데이터 형식이 동일해야 한다**. 두 컬럼이 데이터 형식이 동일하고 같은 이름이면 Spotfire는 로드할 때 두 컬럼을 자동으로 일치시킨다. <u>데이터 테이블 속성 대화 상자</u>(뒷 부분에서 상세 설명)에서 컬럼 일치를 보고 편집할 수 있다. 동일한 시각화에서 여러 데이터 테이블을 사용하는 방법에 대한 자세한 내용은 한 시각화의 여러 데이터 테이블을 참조하면 된다.

한 시각화의 여러 데이터 테이블

 Spotfire에서는 서로 다른 두개 이상의 여러 데이터 테이블에 있는 컬럼값들을 하나의 시각화에서 사용할 수 있다. 두 컬럼이 데이터 형식이 동일하고 같은 이름이면
Spotfire는 로드할 때 두 컬럼을 자동으로 일치시킨다. 아래 예제에서는 별도로 두 테이블을 Join하지 않아도 자동으로 일치시켜서 시각화(차트)를 구성하고 있다.



Sales 2011

Category	Туре	Sales 2011
Fruit	Apples	19000
Fruit	Pears	22000
Fruit	Bananas	29000
Vegetables	Cucumber	12000
Vegetables	Tomatoes	25000
Vegetables	Lettuce	19000

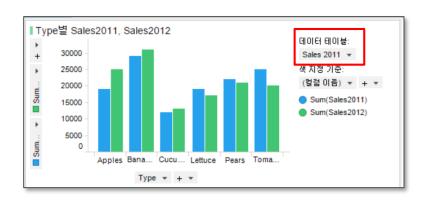


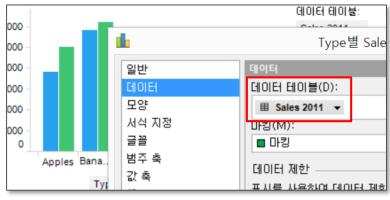
Sales 2012

Category	Туре	Sales 2012
Fruit	Apples	25000
Fruit	Pears	21000
Fruit	Bananas	31000
Vegetables	Cucumber	13000
Vegetables	Tomatoes	20000
Vegetables	Lettuce	17000

기본 데이터 테이블

- 여러 데이터 테이블의 데이터를 결합하는 시각화에서 기본 데이터 테이블은 중요한 역할을 수행한다. 시각화에서 기본 데이터 테이블은 항상 하나만 있으며 시각화의 데이터에 있는 앵커 지점이다.
- 시각화에서 항목을 마크할 경우 기본 데이터 테이블의 컬럼에 대한 상세 정보만 표시됩니다. 기본적으로 모든 표현식은 기본 데이터 테이블을 참조한다.
- 데이터를 최대한 활용하기 위해서는 시각화 구성을 시작하기 전에 어떤 데이터 테이블이 시각화의 기본 데이터 테이블로 가장 적합한지 생각해 보는 것이 좋다.
- 시각화 속성 대화 상자의 데이터 페이지나 범례의 데이터 테이블 선택기에서 기본 데이터 테이블을 선택할 수 있다. 아래 이미지에서 기본 데이터 테이블의 이름은 'Sales 2011'이다.





추가 데이터 테이블

- 기본 데이터 테이블이 아닌 데이터 테이블의 컬럼은 시각화의 집계 축에서는 사용할 수 있지만 시각화를 그룹화하는 축에서는 사용할 수 없다.
- 다른 데이터 테이블의 컬럼을 추가하려는 경우 데이터 패널이나 필터 패널에서 드래그 앤 드롭을 사용하거나 컬럼 선택기에서 컬럼을 선택할 수 있다. 컬럼 선택기를 열고 원하는 데이터 테이블로 전환하면 된다. 컬럼 선택기가 전환되어 선택한 데이터 테이블의 컬럼이 대신 표시된다. 데이터 테이블 선택기는 분석에 여러 데이터 테이블이 있고 사용 가능한 컬럼 일치가 있는 경우에만 표시된다(아래 참조).

50000 Sales 2011 -☼ ∏+ Sales 2011 Sales 2012 Edit Column Matches... Category Type Sales 2011

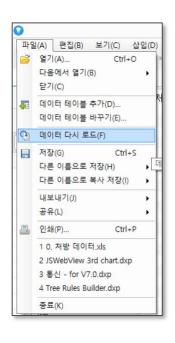
컬럼선택기

■ 컬럼 일치

- 알고 있어야 하는 또 하나의 중요한 개념은 컬럼 일치이다. 시각화에서 여러 데이터 테이블의 데이터를 표시하려면 시각화를 일정 방식으로 그룹화하는 데 사용할, 하나 이상의 컬럼이 시각화의 다른 데이터 테이블에 있는 해당하는 컬럼과 일치해야 한다. 컬럼이 일치하려면 컬럼에 동일한 종류의 데이터가 포함되어 있어야 한다. 컬럼이 동일한 데이터 형식의 값을 포함하고 이름이 같으면 자동으로 일치됩니다. 예를 들어 아래의 두 데이터 테이블 'Sales 2011' 및 'Sales 2012'에서 컬럼 'Category'와 'Type'은 두 데이터 테이블 간에 일치된다.
- 시각화를 설정할 때 기본 규칙은 시각화에서 사용할 모든 범주가 모든 데이터 테이블에 있어야 한다는 것이다.

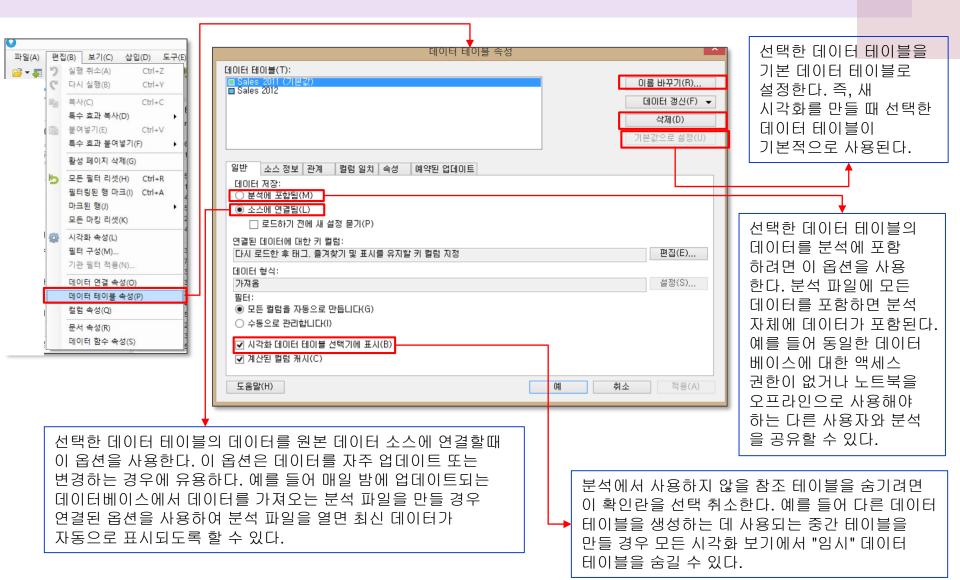
Sales 2011			S	Sales 2012		
Category	Туре	Sales 2011	Category	Туре	Sales 2012	
Fruit	Apples	19000	Fruit	Apples	25000	
Fruit	Pears	22000	Fruit	Pears	21000	
Fruit	Bananas	29000	Fruit	Bananas	31000	
Vegetables	Cucumber	12000	Vegetables	Cucumber	13000	
Vegetables	Tomatoes	25000	Vegetables	Tomatoes	20000	
Vegetables	Lettuce	19000	Vegetables	Lettuce	17000	

1. 데이터 기본 : 데이터테이블 다시로드

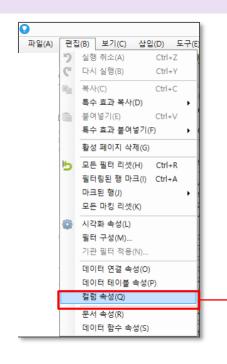


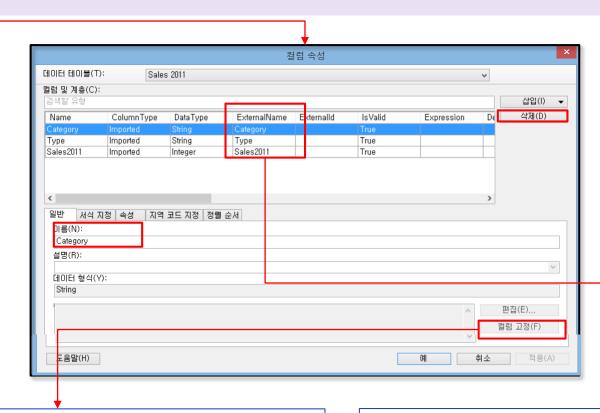
- Spotfire에서는 분석 결과를 저장(.dxp)할 때 몇 가지 옵션이 있는데, 데이터를 .dxp 파일 내부에 포함하여 저장할 것인지, 아니면 매번 분석 파일(.dxp)을 불러올 때마다 원래 작성시 참조했던 외부 source를 새로 참조하여 계산을 할지 선택할 수가 있다.
- 만일 후자의 경우로 저장 옵션이 선택 되어 있는 경우, 원래 source의 값이 변경되어 있다면, Spotfire를 사용하는 중간에 언제든지 새로운 값을 반영하여 분석 결과를 update할수가 있다.
- 예를 들어, xls에 있는 값을 가져와서 분석하였는데, xls에 있는 일부 컬럼의 값이 변경되어서 변경된 부분이 Spotfire에서 바로 반영되도록 하고 싶을 때 '데이터 다시 로드'를 실행하면 Spotfire 는 변경된 파일의 값으로 즉시 updat된다.
- 이는 Database에서 데이터 테이블을 참조하여 작성된 경우에도 동일하게 적용된다.
- 주의: 예를 들어 xls의 경우, 값을 수정했다면 반드시 먼저 xls에서 저장하기를 누른 후에 Spotfire에서 '데이터 다시 로드'를 실행해야 한다.

1. 데이터 기본 : 데이터테이블 속성



1. 데이터 기본 : 컬럼 속성



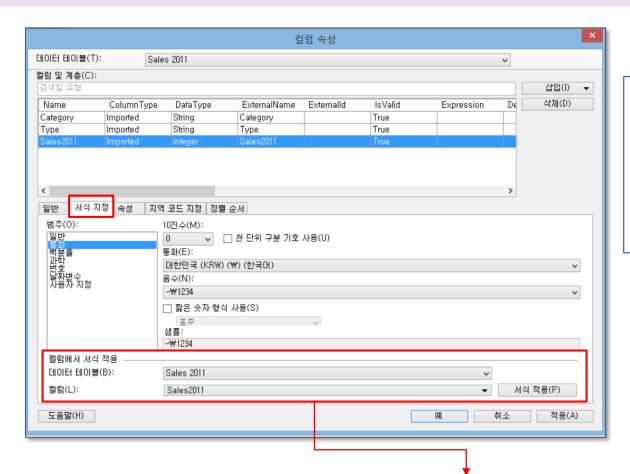


컬럼을 더 이상 편집할 수 없도록 컬럼을 잠근다. 그러면 이 컬럼을 다시계산하기 위해 데이터 바꾸기 작업이 수행되지 않는다. 계산됨, 함, 결과, 태그 및 마스크(필터 포함/필터 제외) 형식 컬럼을 고정할 수 있다. 가져온일반 컬럼은 고정하거나 계층화 할 수 없다.

덮어쓰지 않도록 하거나 성능상의 이유로 계산 결과를 저장하려는 경우에 고정할 수 있다. 데이터 테이블에 컬럼을 추가할 때 조인 조건에서 컬럼을 사용할 수 있도록 하려는 경우에도 컬럼을 고정할 수 있다. 최초에 데이터 테이블을 가져올 때의 original name.

즉, 사용자가 Spotfire에서 컬럼 이름을 변경하였을 경우에, DB나 XIs에서 원래 가져올 때 최초에 불러왔던 원본 source에 저장되어 있는 컬럼 이름을 표시한다.

1. 데이터 기본 : 컬럼 속성(서식 지정)



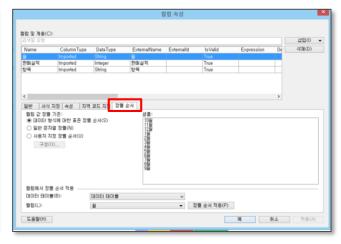
- 컬럼 수준에서 값을 서식(format) 지정할 수 있다.
- 이 대화 상자에서 특정 컬럼 또는 계층에 대한 설정을 변경한 경우 이후의 모든 분석에서 해당 컬럼 또는 계층에 새 설정이 사용된다.

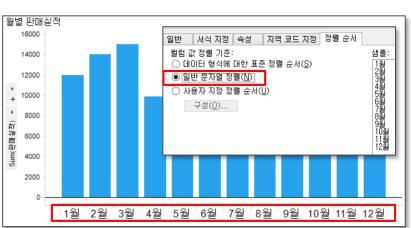
• 원하는 컬럼의 서식을 위의 '서식 지정' 탭에서 지정하지 않고, 이미 지정해 놓은 다른 테이블이나 컬럼의 서식을 참조하여 그 설정을 똑 같이 적용하고 싶을 때 사용한다.

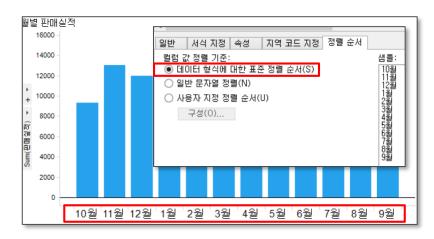
1. 데이터 기본 : 컬럼 속성(정렬 순서)

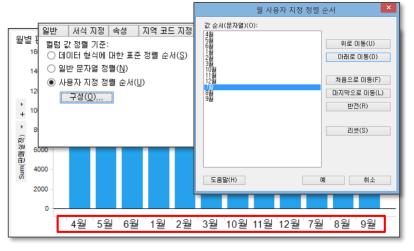
 선택한 컬럼이 문자(String) 형식인 경우, 정렬 순서를 데이터 형식의 표준 정렬 순서에서 일반 문자열 정렬 또는 사용자 지정 정렬 순서로 변경할 수 있다



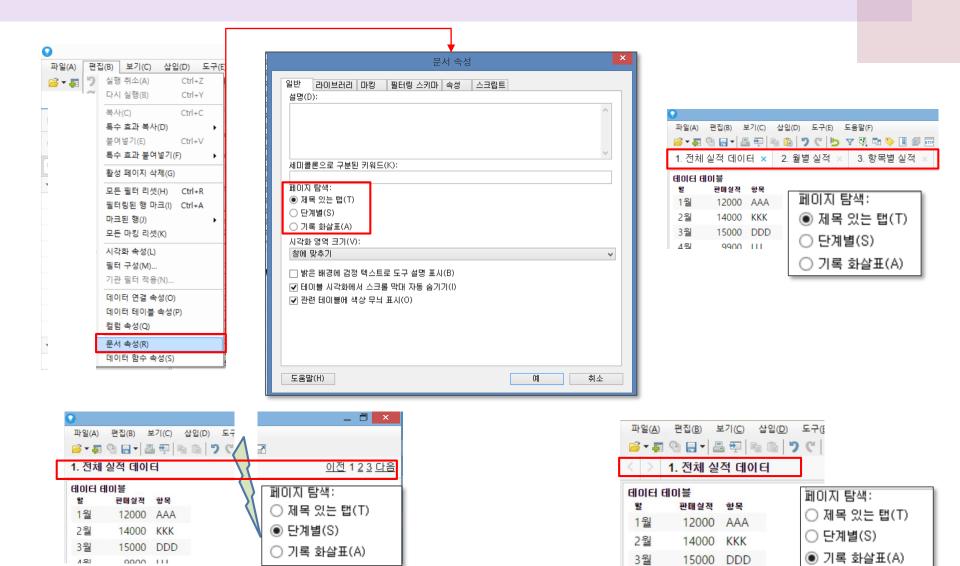




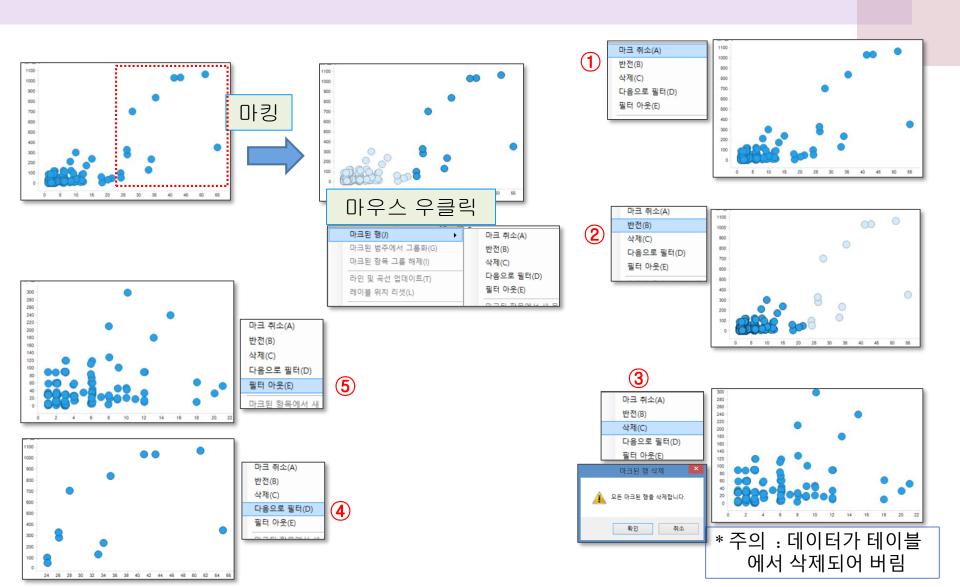




1. 데이터 기본 : 문서 속성(일반)

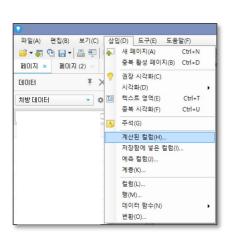


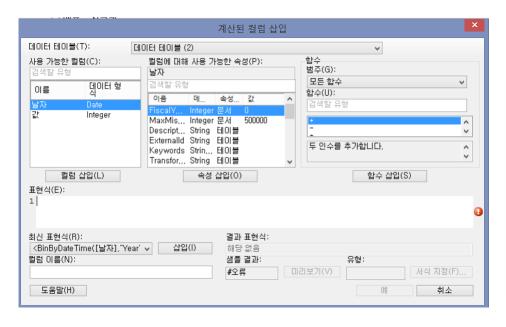
1. 데이터 기본 : 마킹 데이터 핸들링



1. 데이터 기본 : 삽입 > 계산된(Calculated) 컬럼

- 현재 Spotfire 에 load된 데이터 테이블에 포함된 컬럼만으로 필요한 연산을 수행할 수 없거나 필요한 시각화를 만들 수 없는 경우가 종종 있다. 이 때 기존 컬럼들 값들간의 산술 및 논리 표현식을 사용하여 원하는 새로운 컬럼을 만들어서 사용할 수 있다.
- 참고: 계산된 컬럼은 다른 컬럼과 마찬가지로 취급되며 그 컨텐츠는 그 후의 모든 분석 과정에서 정적인 상태이다.

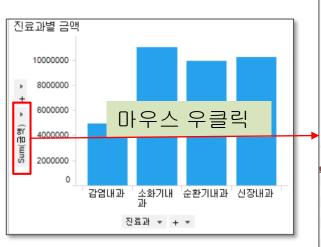


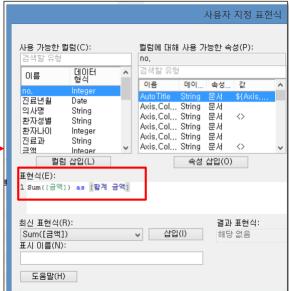


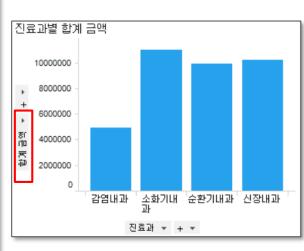


1. 데이터 기본 : 삽입 > 계산된(Calculated) 컬럼

- 컬럼 이름을 "[" 및 "]"(꺽쇠괄호)로 묶어 컬럼에 액세스할 수 있다.
- 변수, 함수 및 키워드는 대/소문자를 구분하지 않는다. SUM(C1) = Sum(C1) = sum(C1)
- 새로 만들어진 컬럼은 데이터 테이블의 나머지 모든 컬럼과 행 수가 동일하다.
- 여러 개의 컬럼은 기본적으로 쉼표로 구분된다.
- 여러 표현식을 사용하는 경우에는 AS 키워드를 사용하여 사용자 지정 표현식 대화 상자에서 표현식 이름을 변경할 수 있다.

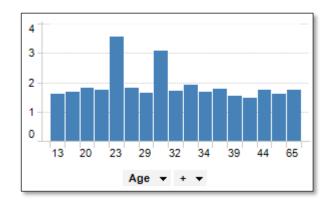


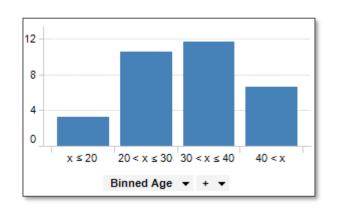




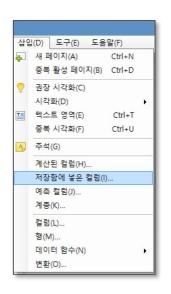
1. 데이터 기본 : 삽입 > 저장함에 넣은 컬럼

- 저장함을 사용하면 많은 수의 연속하는 값을 적은 수의 "저장함"으로 그룹화할 수 있습니다. 예를 들어 일련의 사람에 대한 데이터가 있는 경우 연령을 적은 수의 연령 간격으로 정렬할 수 있습니다. 컬럼 선택기를 마우스 오른쪽 단추로 클릭하고 저장함에 컬럼 자동 저장을 클릭하여 숫자 컬럼을 일시적으로 그룹화할 수도 있습니다.
- 범주별 값을 저장함에 그룹화하는 옵션도 있습니다. 이는 컬럼의 범주별 값이 필요 이상으로 많은 경우에 유용합니다. 예를 들어 사과, 배, 오렌지 및 라임의 판매량을 보여주는 시각화에서 감귤류 판매량과 사과 및 배 판매량을 비교해볼 수 있습니다. 그리고 오렌지와 라임을 저장함에 그룹화할 수 있습니다.



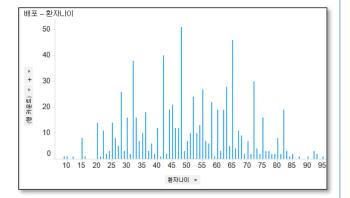


1. 데이터 기본 : 삽입 > 저장함에 넣은 컬럼

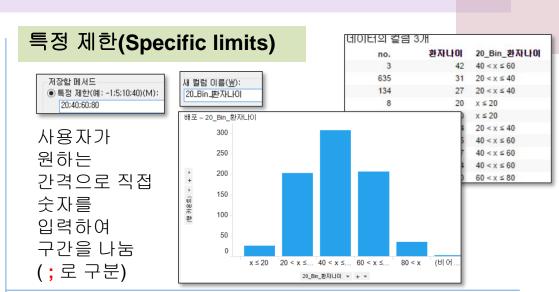


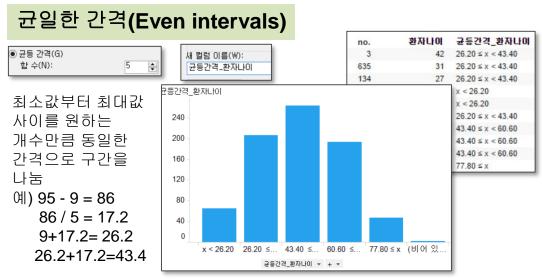
한
γH
환자나이
42
31
27
20
20
34
55
47
44
80

원본 데이터 테이블



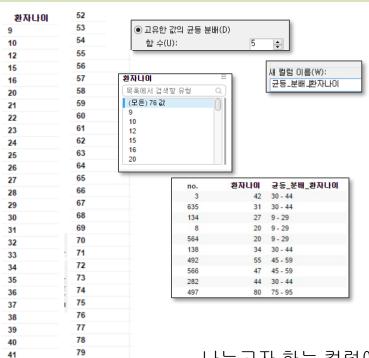
Original '환자 나이 ' 컬럼





1. 데이터 기본 : 삽입 > 저장함에 넣은 컬럼

고유한 값의 균등 분배(Even distribution of unique values)



42

43

45

48

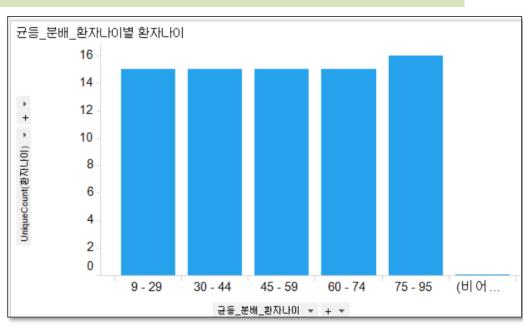
81

82

87

90

(비어 있음)



나누고자 하는 컬럼에 있는 모든 고유한 값들(uniqe value) 사이를 원하는 개수만큼 동일한 간격으로 구간을 나눔

예) '환자나이' 컬럼에는 9,10,12,15, ... 93, 95까지 총 75개의 unique value가 존재함 → 95 / 5 = 15(15개씩으로 한 구간을 나누면 됨) → 9 ~ 29 사이에 총 15개, 30~44 사이에 총 15개 ... 75~95 (마지막 16개) 이런 식으로 구간을 나눔.

1. 데이터 기본 : 삽입 > 계층

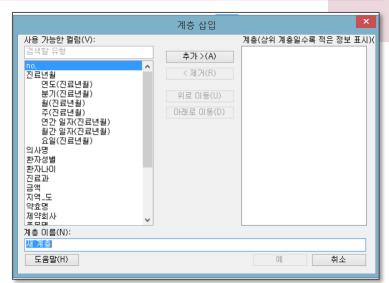
기존 컬럼들을 이용하여 컬럼들간의 계층 구조를 갖는 컬럼을 만들고 이것을 '필터'로서도 이용할 수가 있다.

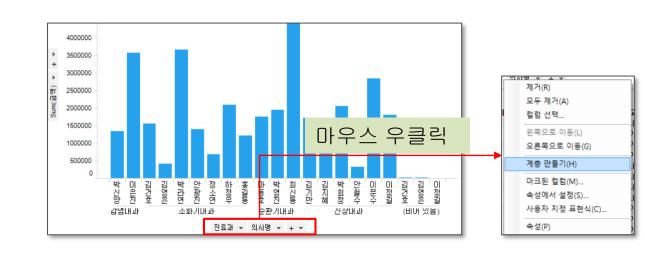
예를 들어, 대분류>중분류>소분류 의 경우가 해당된다.

계층 구조의 컬럼을 만드는 방법은 두 가지가 있다.

- 1. 메뉴의 '삽입' > '계층'
- 시각화에서 계층 구조를 갖는 컬럼을 축으로 2개 이상 선택한 후 마우스 우클릭하여 작성

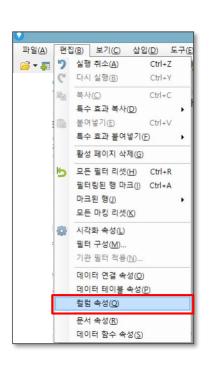


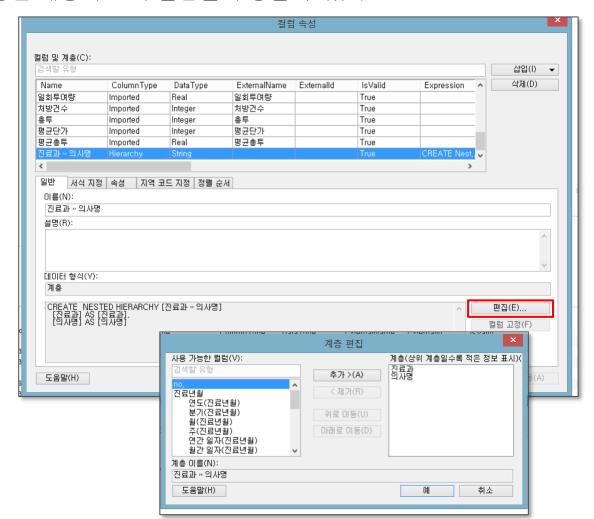




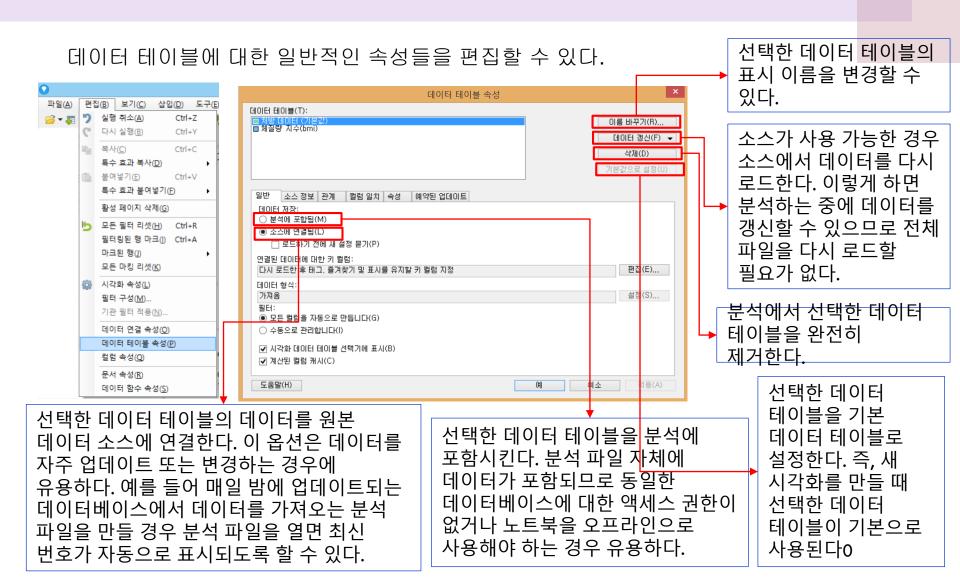
1. 데이터 기본 : 계층 구조 편집(수정)

사용자가 만들어 놓은 계층 구조의 컬럼을 수정할 수 있다.

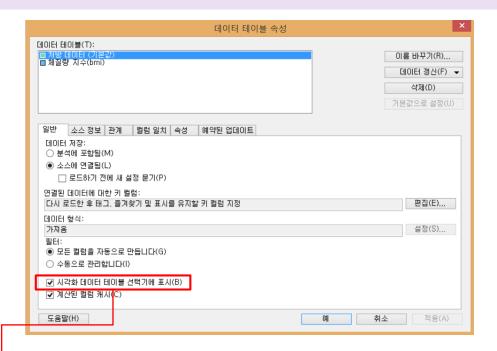




1. 데이터 기본 : 데이터 테이블 속성



1. 데이터 기본 : 데이터 테이블 속성

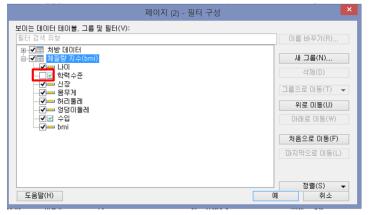


분석에서 자주 사용하지 않는, 참조용 테이블을 숨길때면 이 확인란의 선택을 취소한다. 예를 들어 다른 데이터 테이블을 생성하는 데 중간 단계로서 생성한 테이블을 만들었을 경우, 모든 시각화 보기에서 "임시" 데이터 테이블을 숨길 수 있다. 이후 모든 시각화 관련 정보 창에서 '시각화 데이터 테이블 선택기에 표시'를 해제한 테이블을 선택할 수 없도록 됨을 확인할 수 있다.



마우스를 필터 부분에 놓고 마우스 우클릭을 하여 '필터 구성'을 이용하여 원하는 필터 구성을 할 수 있다.





For Join 디렉토리/ '고객정보.xls' + '판매정보.xls'

2. 데이터 연결(Join)

서로 다른 2개의 데이터 테이블을, 동일한 값들을 갖는 컬럼(Join Key)을 이용하여 서로 연결시키는 방법이다.

D	이름	성별	LHOI	지역	최초 회원 가	데이터 테이블:
ond022	김도빈	남	38	서울	2008-02-10	고객정보 🔻
ond101	박진수	남	145	경기	2006-11-16	
ond278	최윤희	q	24	강원	2006-12-15	
ond171	이기영	남	33	전남	2007-01-04	
ond050	김지수	여	28	전북	2006-11-18	
ond196	이원균	남	40	경남	2006-12-20	
ond018	권진경	여	36	경북	2007-01-05	
ond148	유인나	여	50	충남	2007-01-02	
ond096	박정일	남	46	중북	2006-11-18	
ond141	모햇살	여	31	제주	2006-12-31	
ond133	안승윤	남	20	대전	2007-01-06	
ond162	윤진건	남	33	대구	2007-01-12	
1400	II A W	1.6		н.п	0000 00 40	



위 그림 좌측에 '고객정보 '테이블과 우측에 '판매 실적'데이터 테이블이 있다. 현재는 아직 연결이 되지 않았기 때문에 각 화면 우측에 줄무늬 색 상이 각각 녹색과 파란색으로 서로 다름을 알 수 있다.

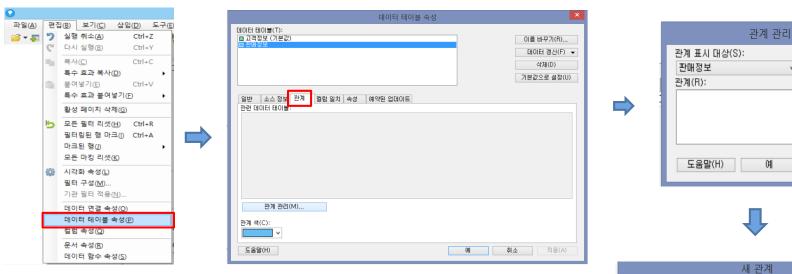
신규(N)...

취소

For Join 디렉토리/ '고객정보.xls' + '판매정보.xls'

2. 데이터 연결(Join)

2개의 데이터 테이블에서 Join Key에 해당하는 컬럼을 각각 지정하여 Join(관계)를 설정시킨다.



우측 그림에서 이제 두개의 서로 다른 데이터 테이블이 'ID'란ㄴ 컬럼으로 서로 연결되었음을 표시하고 있다.

관계 관리	X
관계 표시 대상(<u>S</u>): 판매정보	v
관계(<u>R</u>): 고객정보 판매정보 ID ID	신규(Ŋ)
	편집(E) 삭제(D)
도움말(<u>H</u>) 예	취소

			X
	새 된	관계	
왼쪽 데이터 테이블(<u>L</u>): 고객정보	V	오른쪽 데이터 테이블(<u>R</u>): 판매정보	V
왼쪽 컬럼(<u>C</u>): ID	7	오른쪽 컬럼(<u>O</u>): ID	-
왼쪽 메서드(<u>M</u>): (없음)	¥	오른쪽 메서드(<u>T</u>): (없음)	Y
샘플 값: pnd022		샘플 값: pnd022	
도움말(<u>H</u>)		예 취소	

For Join 디렉토리/ '고객정보.xls' + '판매정보.xls'

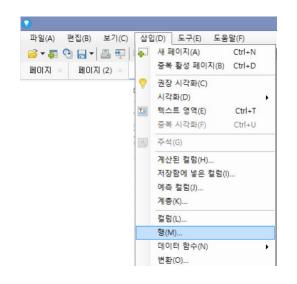
2. 데이터 연결(Join)

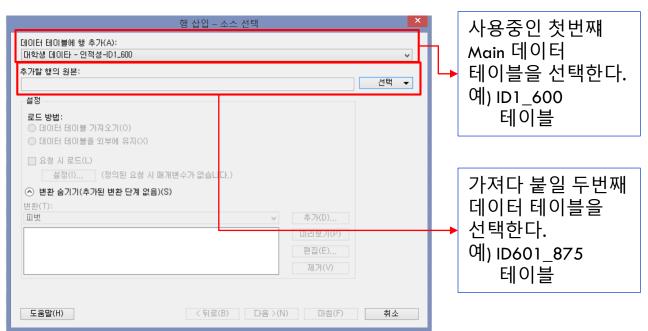


- 이제 2개의 테이블이 하나의 테이블처럼 연결되어 있다.
- 이는 우측의 2개의 테이블에서 각 테이블명을 표시하고 있는 시각화 Title의 좌측 상단에 줄무늬가 이제 첫번째 테이블(고객정보)의 색상으로 두번째 테이블도 변경되었음 으로부터 알 수 있다. 즉, 2개의 테이블이 동일한 줄무늬를 사용하고 있다.
- 연결 키(Join Key)는 원하는 만큼 더 생성하여 사용할 수 있다.

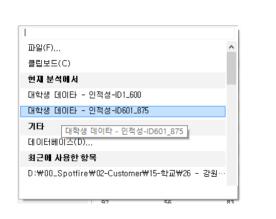
3. 행 삽입 (Add Rows)

- 사용중인 하나의(첫번째) 데이터 테이블에 다른 두번째 데이터 테이블을 행으로(밑으로) 추가할 수 있다. 예를 들어서 학생들의 ID가 1~600번인 'ID1_600' 테이블로 만든시각화를 사용하다가 ID가 601~875번인 'ID601_875' 라는 새로운 데이터를가져와서(데이터 테이블 추가) 현재의 'ID1_600' 데이터에 'ID601_875' 데이터를추가하여 하나의 데이터 테이블로 사용할 수 있다.
- 행 추가 결과, 원래 행의 수가 600개에서 875개로 늘어나게 된다.

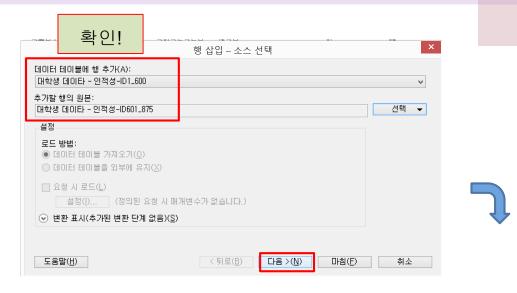


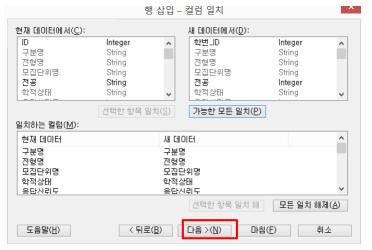


3. 행 삽입 (Add Rows)









만일 두테이블의 컬럼명과 컬럼 속성이 같다면 '가능한 모든 일치'를 눌러서 자동으로 컬럼 일치를 시킨다.

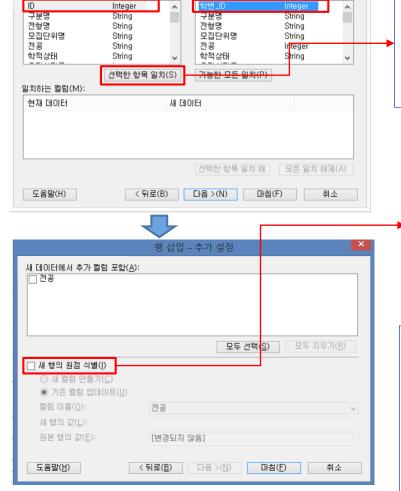


3. 행 삽입 (Add Rows)

행 삽입 _ 컬럼 일치

새 데이터에서(D):

현재 데이터에서(C):



하지만, 좌측 그림과 같이 두개의 컬럼명의 서로 다른 경우(하지만, 반드시 컬럼 속성은 일치해야 한다)에는 현재 데이터와 새 데이터에서 각각 컬럼명을 선택한 후에 '선택한 항목 일치' 버튼을 눌러서 수동으로 연결키를 설정해 주어야만 한다.

만일 현재 데이터(첫번째 데이터)의 컬럼에는 없는데 두번째 새로운 데이터에만 컬럼이 존재하는 경우, 좌측 그림과 같이 남아있는 컬럼(이 경우에는 '전공')이 표시된다.

이 컬럼(들)을 최종 데이터에 추가하고 싶으면 컬럼 앞의 체크박스를 선택한다.

행 삽입시에 첫번째 원본 데이터테이블과 두번째 새로운 데이터테이블을 구분하고자 할 때 이 체크박스를 선택한다.

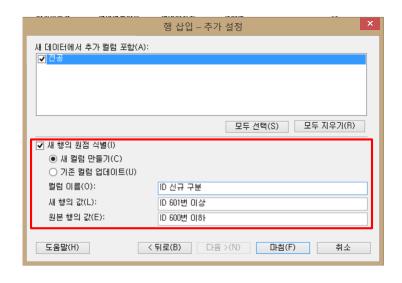
예를 들어, 원본 테이블에 1월이라는 날짜 정보의 컬럼이 존재하지 않는데, 2월이라는 새 데이터를 추가할 때, 기존 데이터는 1월, 새 데이터는 2월 이라고 구분하고자 할때 사용할 수 있다.

3. 행 삽입 (Add Rows)

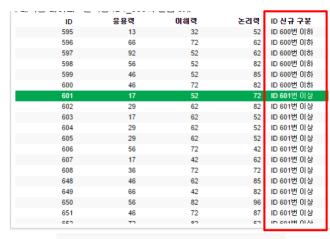


- 행 삽입 이전의 데이터
- 행 수가 600개이며, 원래 초기의 컬럼을 유지하고 있다.

600개 행 중 600개 0개 마크됨







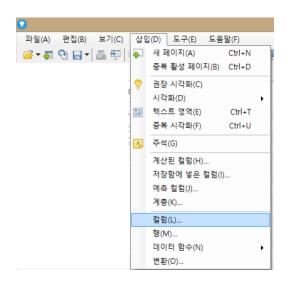
875개 행 중 875개 0개 마크됨

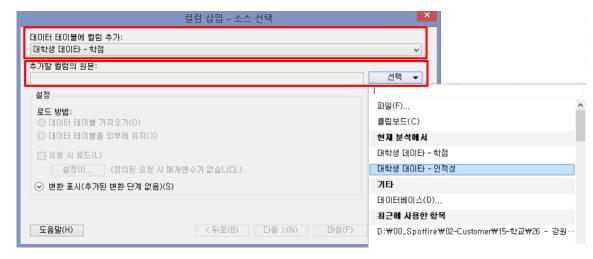
- 행 삽입 이후의 데이터
- 행 수가 875개로 늘어 났으며, 행 삽입시 새로 추가한 컬럼(ID 신규 구분)도 보인다.

Add Column 디렉토리/ 'xxx - 학점.xls' + 'xxx -인적성.xls'

4. 컬럼 추가 (Add Columns)

- 사용중인 하나의(첫번째) 데이터 테이블에 다른 두번째 데이터 테이블을 컬럼으로 (옆으로) 추가할 수 있다. 예를 들어서 학생들의 학점 정보가 있는 데이터 테이블로 만든 시각화를 사용하다가 해당 학생들의 인적성과 관련된 정보가 들어있는 새로운 데이터를 가져와서(데이터 테이블 추가) 현재의 학점 데이터에 인적성 데이터를 추가하여 하나의 데이터 테이블로 사용할 수 있다.
- 컬럼 추가 결과, 원래 컬럼의 수보다 더 많은 컬럼으로 늘어나게 된다.





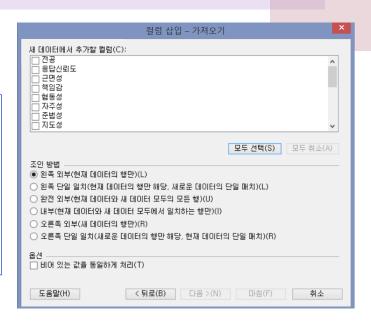
Add Column 디렉토리/ 'xxx - 학점.xls' + 'xxx -인적성.xls'

4. 컬럼 추가 (Add Columns)





만일 두테이블의 컬럼명과 컬럼 속성이 같다면 '가능한 모든 일치'를 눌러서 자동으로 컬럼 일치를 시킨다.

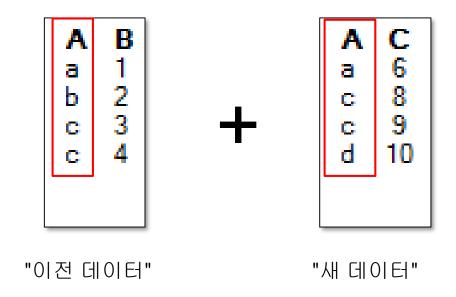


하지만, 만일 두개의 컬럼명의 서로 다른 경우(하지만, 반드시 컬럼 속성은 일치해야 한다)에는 현재 데이터와 새 데이터에서 각각 컬럼명을 선택한 후에 '선택한 항목 일치' 버튼을 눌러서 수동으로 연결키를 설정해 주어야만 한다.

'행 추가'의 경우와는 달리, 컬럼 추가시에는 연결하는 방법(Join Method)에 대하여 신중하게 고려를 해야만 한다. 연결하는 방법에는 총 6가지의 옵션이 존재한다.

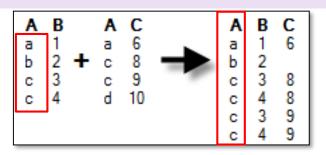
4. 컬럼 추가 (Add Columns)

원래 있던 데이터 테이블에 새 데이터 테이블의 컬럼을 추가하는 방법



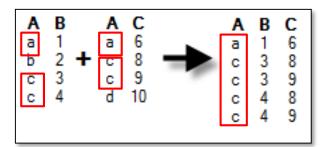
두 데이터 테이블이, 두 데이터 테이블의 "A" 컬럼에 있는 지시자(Join Key)를 사용하여 일치된다.

4. 컬럼 추가 (Add Columns)



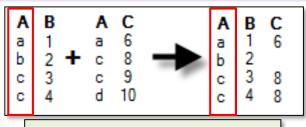
왼쪽 외부 (Left outer)

왼쪽 테이블을 기준 으로 가능한 모든 조인



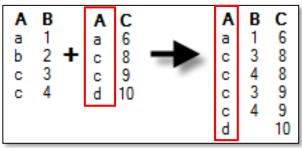
내부 (Inner)

둘 다 값이 일치하는 경우만 존재. 한쪽에 값이 없으면 삭제



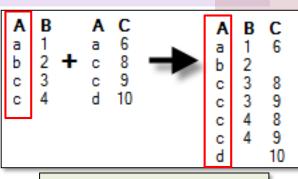
왼쪽 단일 일치 (Left single match)

왼쪽 테이블의 행의 수가 늘지 않음. 첫번째 만나는 값으로 결정



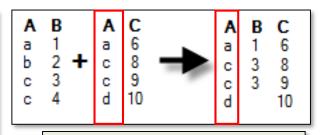
오른쪽 외부 (Right outer)

오른쪽 테이블을 기준으로 가능한 모든 조인



완전 외부 (Full outer)

둘 다 모든 경우 다 존재. 값이 없으면 빈칸으로 존재.



오른쪽 단일 일치 (Right single match)

오른쪽 테이블의 행의 수가 늘지 않음. 첫번째 만나는 값으로 결정