

중복되는 국가의 이름 타이디한 데이터 세트 만들기

변수 설명을 보면 country는 국가 이름을 의미하고 iso2와 iso3는 국가 코드를 2자나 3자로 줄인 열 이다. 이는 국가를 중복해서 지정하는 변수가 세 개임을 의미한다. 따라서 iso2와 iso3는 중복되고, 타이디한 데이터 세트를 만들기 위해서는 삭제해야 한다.

원래의 데이터 세트는 아래와 같은 모습을 보인다.

	country	iso2	iso3	year	new_sp_m... ¹	new_s... ²	new_s... ³	new_s... ⁴	new_s... ⁵	new_s... ⁶	new_s... ⁷	new_s... ⁸	new_s... ⁹	new_s... ¹⁰
	<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>
1	Afghanistan	AF	AFG	1980	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	Afghanistan	AF	AFG	1981	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	Afghanistan	AF	AFG	1982	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
4	Afghanistan	AF	AFG	1983	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	Afghanistan	AF	AFG	1984	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	Afghanistan	AF	AFG	1985	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	Afghanistan	AF	AFG	1986	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	Afghanistan	AF	AFG	1987	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	Afghanistan	AF	AFG	1988	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10	Afghanistan	AF	AFG	1989	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

country는 나라 이름을 나타내고 iso2는 나라 이름의 맨 처음 앞글자 두 개를 값으로 사용하고 있다. iso3는 나라 이름의 맨 처음 앞글자 세 개를 값으로 이용하는 모습을 보인다. iso2, iso3는 country와 중복되는 열을 의미하므로 삭제해야 한다.

원래의 데이터 세트는 타이디 하지 않은 모습을 보이고 있고 이를 타이디 데이터 세트로 최종적으로 가공하면 아래와 같은 데이터 세트를 가지게 된다.

```
*.☆
+;:☆* 중복 삭제!
"+.;* | ^ ^
(. * * n · ω · )
☆:+. ;, | , |
+*: * +*'' じ, — J
*:+. ☆
```

	country	year	type	sex	age	cases
	<chr>	<int>	<chr>	<chr>	<chr>	<int>
1	Afghanistan	1997	sp	m	014	0
2	Afghanistan	1998	sp	m	014	30
3	Afghanistan	1999	sp	m	014	8
4	Afghanistan	2000	sp	m	014	52
5	Afghanistan	2001	sp	m	014	129
6	Afghanistan	2002	sp	m	014	90
7	Afghanistan	2003	sp	m	014	127
8	Afghanistan	2004	sp	m	014	139
9	Afghanistan	2005	sp	m	014	151
10	Afghanistan	2006	sp	m	014	193

타이디한 데이터 세트는 iso2, iso3가 지워진 모습을 볼 수 있다.

"Written by Choi, Ji-Hyeon"

∠(↯↯)∠_