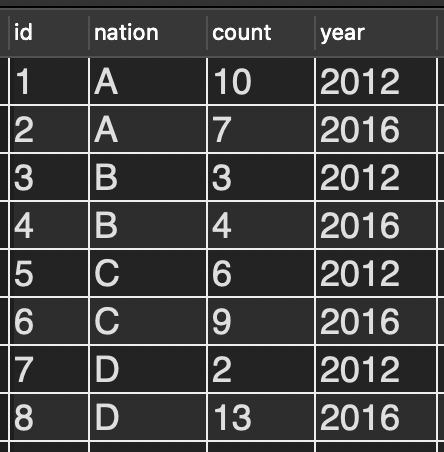
이전 영상 후반에서 두 테이블의 합집합을 구해주는 집합 연산자인 **UNION**을 배웠습니다. UNION에 관해서 알아두면 좋은 2가지 사실이 있는데요. 하나씩 살펴볼게요.

**1. 서로 다른 종류의 테이블도, 조회하는 컬럼을 일치시키면 집합 연산이 가능합니다.**

다음과 같은 두 테이블이 있습니다.

**Summer\_Olympic\_Medal** : 국가별 하계 올림픽 메달 수 테이블

**Winter\_Olympic\_Medal** : 국가별 동계 올림픽 메달 수 테이블



**Summer\_Olympic\_Medal 테이블**

이 테이블에는 **id(Primary Key), nation(국가), count(메달 수), year(올림픽 개최 연도)** 컬럼이 있습니다.



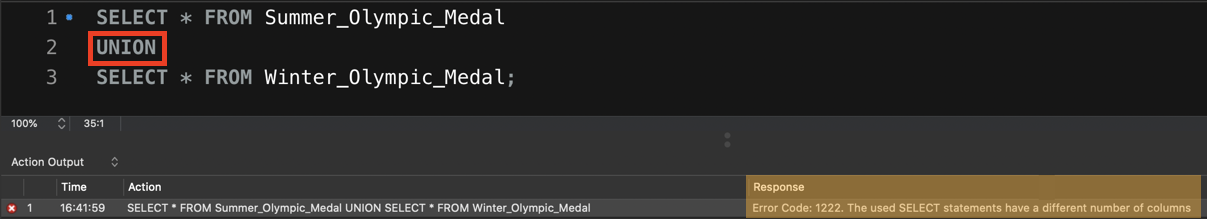
**Winter\_Olympic\_Medal 테이블**

이 테이블에는 **id(Primary Key), nation(국가), count(메달 수), location(올림픽 개최 도시), first\_rank\_count(메달 획득 1위 국가의 메달 수)** 컬럼이 있습니다.

저는 두 테이블을 **UNION 연산**해서 각 국가의 메달 획득 수를 한 눈에 보고 싶습니다.

하지만 이전에 배웠던 것처럼 컬럼 구조가 같은 테이블끼리만 UNION 연산을 할 수 있는데요.

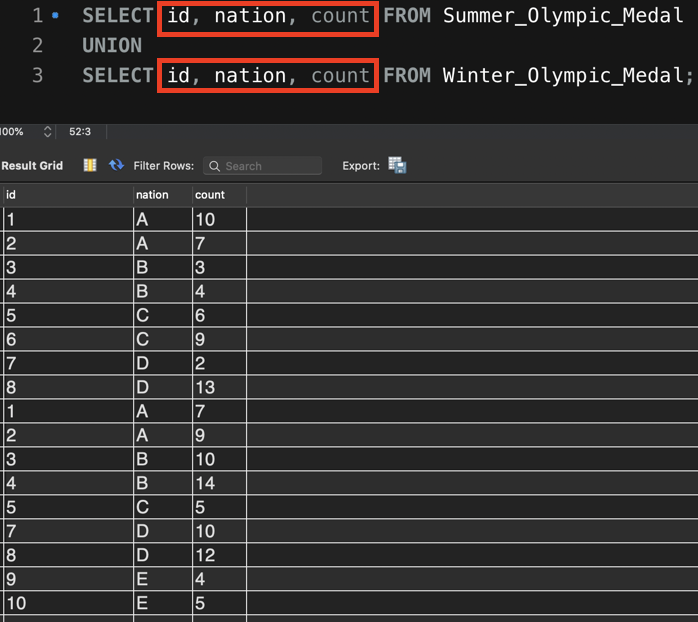
지금 상태에서 바로 UNION 연산을 하면



이렇게 에러가 나는 것을 볼 수 있습니다. 에러 메시지를 보면 두 테이블의 컬럼 수가 다르다는 메시지가 보입니다. 컬럼 구조가 달라서 UNION 연산이 실패한 거죠.

하지만 **방법이 있습니다. SELECT 절 뒤의 \* 부분을 두 테이블이 공통적으로 갖고 있는 컬럼 이름들로 바꿔주면 됩니다.**

위 그림에서 \*(별표) 부분들을 두 테이블의 공통 컬럼들인 **id, nation, count**로 바꿔주고 실행하면,



이렇게 두 테이블의 해당 컬럼들만을 대상으로 UNION 연산이 된 것을 볼 수 있습니다. 방금 본 것처럼 두 테이블의 원래 컬럼 구조가 달라도, 두 테이블이 공통적으로 갖고 있는 컬럼들만 조회한 경우에는 UNION 같은 집합 연산을 수행할 수 있다는 사실, 잘 기억하세요.

(총 컬럼의 수와, 각 컬럼의 데이터 타입만 일치하면 UNION 연산이 가능합니다)

**2. UNION 과 UNION ALL**

그런데 합집합을 구해주는 UNION에는 특징이 하나 있습니다.

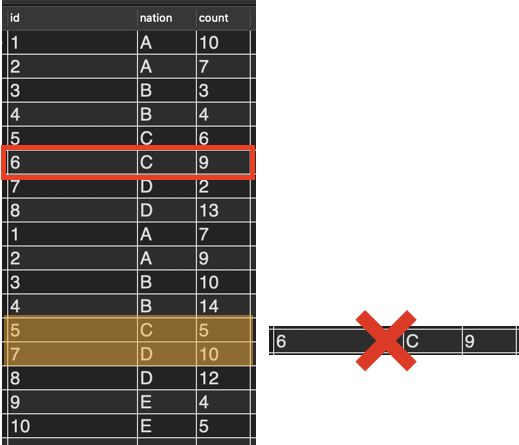
그건 바로, UNION은 두 테이블이 공통적으로 갖고 있는 원소들, 그러니까 두 테이블의 교집합에 해당하는 영역의 row들은 **중복을 제거하고, 그냥 딱 하나의 row만 보여준다**는 겁니다. 이것이 UNION의 특징인데요.

방금 전 두 올림픽 메달 테이블을 잠깐 볼까요?



지금 보니까 C 국가는 2016 하계 올림픽에서 9개의 메달을, 어떤 동계 올림픽에서도 9개의 메달을 획득했네요. 우연히도 획득한 메달 수가 같군요. 이 둘은 별개의 데이터입니다. 하지만 우리는 지금 빨간색 박스 영역까지의 컬럼들만을 조회했기 때문에 UNION은 이 두 row를 서로 같은 것으로 판단합니다.

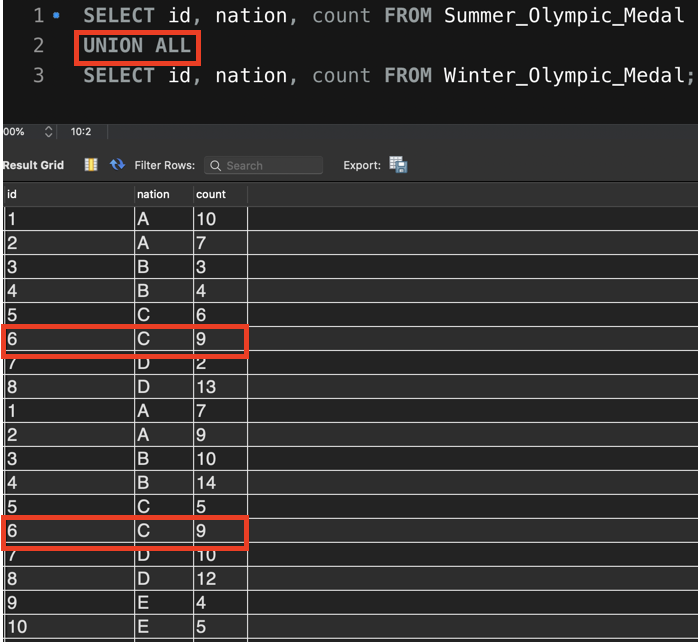
그리고 UNION 연산자는 방금 전 말한대로 중복을 제거하고 하나의 row만 보여줄 겁니다. UNION 연산의 결과를 보면,



겹치는 두 개의 row 중 하나만 표시했다는 것을 알 수 있습니다.

그런데 이렇게 보는 것은 정확한 결과가 아닙니다. 정보가 누락되어 버렸는데요. 이 상황에서는 겹치는 row도 다 그대로 보여주는 것이 맞습니다.

이럴 때는 **UNION** 말고 **UNION ALL**이라는 연산자를 사용하면 됩니다. **UNION ALL은 UNION처럼 두 테이블의 합집합을 보여준다는 점은 같습니다. 하지만 겹치는 것을 중복 제거하지 않고, 겹치는 것들을 그대로 둘다 보여준다는 차이점이 있죠.** 이번에는 UNION 대신 UNION ALL 연산을 해보겠습니다.



결과를 보니 두 테이블 간 겹치는 row들이 둘다 잘 보입니다. 이제 누락되는 정보는 없네요.

방금 본 것처럼 UNION 연산과 UNION ALL 연산은 둘다 합집합을 구하되, 전자는 중복을 제거해서 보여주고, 후자는 그런 작업없이 두 테이블을 합친 결과를 그대로 보여준다는 차이가 있습니다.

만약 중복을 제거하고 깔끔하게 보는 것이 중요한 경우에는 UNION 연산자를 사용하고, 중복을 제거하게 되면 정보 누락이 발생할 수 있는 경우에는 UNION ALL 연산자를 사용하면 되겠죠? 여러분의 상황에 따라 적합한 연산자를 사용하시면 됩니다.