

오, 집계 함수(Aggregation)의 심화 단계까지 오셨군요! 분석 프로젝트를 하다 보면 보고서용 데이터를 뽑을 때 필수적인 기능들이죠.

말씀하신 내용 중에 맞는 부분도 있지만, 살짝 교정이 필요한 부분이 있어요. 특히 GROUPING SETS에 대해 조금 오해가 있으신 것 같아 깔끔하게 정리해 드릴게요.

1. GROUPING SETS: "내가 원하는 조합만 골라잡기"

GROUPING SETS는 칼럼별 집계만 구하는 용도가 아니라, 내가 정의한 모든 조합을 다 구할 수 있는 가장 범용적인 도구입니다. 총계(Grand Total)도 포함할 수 있습니다.

- GROUPING SETS (A, B, ()): A별 집계, B별 집계, 그리고 **전체 총계()**를 한 번에 구합니다.
- 즉, ROLLUP이나 CUBE로 나오는 결과 중 "필요 없는 건 빼고 내가 원하는 조합만" 명시할 때 씁니다.

2. 핵심 3인방 비교 (MySQL/BigQuery 공통 개념)

구분	특징	결과 조합 (A, B 기준)
ROLLUP	계층적 집계 (상사-부하 느낌)	(A, B), (A), ()
CUBE	모든 경우의 수 (다차원 분석)	(A, B), (A), (B), ()
GROUPING SETS	사용자 정의 집계	작성자가 괄호 안에 넣은 것만!

3. UNION ALL과의 관계

사실 ROLLUP, CUBE, GROUPING SETS는 모두 **UNION ALL을 여러 번 쓰는 수고를 덜어주는 '문법적 설탕(Syntactic Sugar)'**입니다.

예를 들어, ROLLUP(A, B)는 아래와 같이 노가다(?)를 하는 것과 결과가 같습니다.

SQL

```
SELECT A, B, SUM(sales) FROM t GROUP BY A, B
UNION ALL
SELECT A, NULL, SUM(sales) FROM t GROUP BY A
UNION ALL
SELECT NULL, NULL, SUM(sales) FROM t -- 총계
```

이걸 GROUP BY ROLLUP(A, B) 한 줄로 끝내는 것이죠.

4. 팩트 체크 및 교정

사용자님: "grouping sets 는 칼럼별합계만을 구할때 사용된다."

Gemini: 반반 맞습니다! 칼럼별 합계(A별, B별)를 구할 때 아주 유용하지만, ()를 추가하면 총계도 뽑을 수 있고, (A, B)를 넣으면 상세 내역도 뽑을 수 있는 아주 유연한 녀석입니다.

실무 팁

지금 데이터 분석 부트캠프에서 배우시는 내용 중, **Tableau** 같은 시각화 툴을 쓰실 때는 SQL에서 굳이 ROLLUP을 안 써도 툴 자체에서 '총계 표시' 버튼 클릭 한 번으로 해결되는 경우가 많습니다.

하지만 **BigQuery**에서 직접 쿼리로 대시보드용 원천 데이터를 만들 때는 ROLLUP을 써서 미리 집계행을 만들어두면 쿼리 횟수를 줄여 비용을 아낄 수 있어요!

혹시 지금 다루고 계신 건강검진 데이터에서 '연령대별 합계'와 '성별 합계'를 동시에 한 표에 나타내고 싶으신가요? 그에 맞는 **GROUPING SETS** 예시 쿼리를 짜 드릴까요?