

sql 독학 강의 # 복수 행(window) 함수 잘 사용 하기(group by) 10편 -sTricky

sTricky 2020. 4. 13. 14:01

sql 독학 강의 # 복수 행(window) 함수 잘 사용 하기(group by) 10편 -sTricky



SQL 독학 강의
복수 행 함수
Group by
사용법

컨텐츠 index

1. group by 절을 이용해 평균 및 합계 구하기
2. group by 절 사용 시 주의할 점

안녕하세요.

지난 시간 복수 행 함수에 관해서 간단한 몇몇 함수들의 예제를 공부해 보았습니다.

이번 포스팅에서는 복수 행 함수를 group by 절을 이용해서 조금 더 세분화하는 내용을 다루어 보도록 하겠습니다.

이번 포스팅에서 사용되는 예제를 다루기 위해서 아래와 같이 테이블을 생성 하고 데이터를 입력하겠습니다.

```
create table class.budget
(
    do varchar(100) null,
    city varchar(100) null,
    budget_value int null,
    population int null
);

INSERT INTO class.budget (do, city, budget_value, population) VALUES ('서울특별시', '서울특별시', 23324, 345);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('부산광역시', '부산광역시', 34323, 5345);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('경상남도', '창원시', 4331, 435);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('경상남도', '양산시', 25436, 2134);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('경상남도', '밀양시', 62341, 6523);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('경기도', '부천시', 3242, 345);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('경기도', '시흥시', 32454, 546);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('경기도', '수원시', 3234, 345);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('충청남도', '공주시', 2425, 436);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('충청남도', '논산시', 5534, 4567);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('강원도', '속초시', 6542, 3542);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('강원도', '강릉시', 23423, 4355);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('강원도', '태백시', 5465, 45);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('전라북도', '전주시', 456, 645);
```

```
);
INSERT INTO class.budget (do, city, budget_value, population) VALUES ('전라북도', '군산시', 3243, 234);
```

위와 같이 데이터를 생성하면 아래와 같이 데이터를 확인할 수 있습니다.

	do	city	budget_value	population
1	서울특별시	서울특별시	23324	345
2	부산광역시	부산광역시	34323	5345
3	경상남도	창원시	4331	435
4	경상남도	양산시	25436	2134
5	경상남도	밀양시	62341	6523
6	경기도	부천시	3242	345
7	경기도	시흥시	32454	546
8	경기도	수원시	3234	345
9	충청남도	공주시	2425	436
10	충청남도	논산시	5534	4567
11	강원도	속초시	6542	3542
12	강원도	강릉시	23423	4355
13	강원도	태백시	5465	45
14	전라북도	전주시	456	645
15	전라북도	군산시	3243	234

각 시별 예산과 인구수를 가상의 데이터로 넣었습니다. 실제와는 큰 차이가 있겠죠?

그럼 이젠 위 데이터를 가지고 group by 절과 함께 어떻게 복수 행 함수를 사용하는지를 함께 공부해 보도록 하겠습니다.

그럼 시작하겠습니다.

#이전 강의 보러 가기#

2020/04/10 - [Database/sql 강의] - sql 독학 강의 # 복수 행(window) 함수 잘 사용 하기(기본 사용법) 9편 -sTricky

sql 독학 강의 # 복수 행(window) 함수 잘 사용 하기(기본 사용...

stricky.tistory.com

1. group by 절을 이용해 평균 및 합계 구하기

위에서 입력한 데이터를 이용해서 각 광역시도별 예산의 평균과 합계를 구해 보도록 하겠습니다.

```
select do, avg(budget_value) as 예산평균, sum(budget_value) as 예산합계
from class.budget
group by do;
```

위 SQL의 결과는 아래와 같습니다.

	do	예산평균	예산합계
1	강원도	11810.0000	35430
2	경기도	12976.6667	38930
3	경상남도	30702.6667	92108
4	부산광역시	34323.0000	34323
5	서울특별시	23324.0000	23324
6	전라북도	1849.5000	3699
7	충청남도	3979.5000	7959

초보자 분들을 위해서 위 내용에 설명을 조금 덧붙이자면, 원 데이터를 먼저 살펴보시면 do라는 컬럼에 "경기도"로 데이터가 들어가 있는 city는 부천시, 시흥시, 수원시가 있습니다. 이 "경기도" 라는 데이터를 한 그룹으로 묶어서 3개의 시에 대한 **예산의 평균 값과, 예산의 합계를 출력한 SQL** 입니다. 이와 마찬가지로 do 라는 컬럼 안에 있는 같은 데이터끼리 묶어서 위와 같이 강원도, 경기도, 경상남도, 부산광역시, 서울특별시, 전라북도, 충청남도의 각각의 예산 평균, 합계 값이 구해져서 나오게 된 것입니다.

이 처럼 group by 절을 사용 하여 이전 포스팅에서 다루었던 모든 복수 행 함수를 위에서 사용한 avg, sum 등과 같이 사용하실 수 있습니다.

2. group by 절 사용 시 주의할 점

group by 절을 사용 할 때 주의하실 점이 몇 개 있습니다. group by 절에는 단순히 컬럼 명을 그대로 써도 좋지만 함수를 이용해서 group by를 할 수도 있습니다. 이때 select 절에도 group by 절에서 쓴 함수 그대로를 써줘야 group by 가 정상적으로 작동한다는 점 이 있습니다.

예를 들어 보겠습니다. 위와 같이 예산 평균, 합계를 구하는데, 수도권과 광역시 두 개의 그룹으로 나누어 보려고 합니다. 편의상 do 컬럼 값 중 "서울특별시", "경기도"를 수도권으로 하고 기타 지역들은 지방으로 나누어 두 그룹의 결과를 아래 SQL을 이용해서 출력해보겠습니다.

```
select do , avg(budget_value) as 예산평균, sum(budget_value) as 예산합계
from class.budget
group by if(do in ('서울특별시', '경기도'), '수도권', '지방');
```

위 SQL의 group by 절을 보면 'if(do in ('서울특별시', '경기도'), '수도권', '지방')' 이렇게 그룹핑을 했습니다. 결과는 아래와 같이 나왔습니다.

※만약 oracle에서 위와 같은 SQL을 실행하면 ORA-00979 가 발생하고 실행되지 않으나, mysql에서는 일단 실행은 됩니다.

	do	예산평균	예산합계
1	서울특별시	15563.5000	62254
2	부산광역시	15774.4545	173519

하지만 group by 절과 select 절에 쓴 내용이 다르기에 결과가 나오긴 했지만 do 칼럼 값이 원하는 대로 나오지 않아 데이터를 활 요하기 힘듭니다.

그래서 아래와 같은 SQL로 수정을 하고 다시 실행을 하여야 합니다.

```
select if(do in ('서울특별시', '경기도'), '수도권', '지방') as 지역구분 , avg(budget_value) as 예산평균, sum(budget_value) as 예산합계
from class.budget
group by if(do in ('서울특별시', '경기도'), '수도권', '지방');
```

결과는 아래와 같이 잘 나왔습니다.

	지역구분	예산평균	예산합계
1	수도권	15563.5000	62254
2	지방	15774.4545	173519

이런 식으로 꼭!! group by 절과 select 절에 그룹핑하는 대상의 형태를 똑같이 작성을 해주셔야 정확한 결과를 가지고 올 수 있다는 것을 명심하여야 합니다.