**[BDA – 8기 SQL 문법 기초반 17주차 필수과제]**

김주경

- 가설을 검증하기 위한 문제에서 제시되는 조건  
1. 연도, 월별 주문 월에 따라 가장 많은 구매 금액을 사용한 나라 Top5 구하기  
2. Top5가 연도별 월별로 동일한 순위를 가지고 있는지 살펴보기

[code – 월별 주문 top 5 확인]

## 1. 월별로 주문 top 5 확인  
with orders\_with\_details as (  
 -- orders 테이블과 orderdetails 테이블을 조인하여 필요한 데이터를 준비한다.  
 select  
 o.orderNumber,  
 o.orderDate,  
 o.customerNumber,  
 c.country,  
 (od.quantityOrdered \* od.priceEach) as total\_amount  
 from orders o  
 join orderdetails od on o.orderNumber = od.orderNumber  
 join customers c on o.customerNumber = c.customerNumber  
 where o.status = 'Shipped'  
),  
monthly\_country\_sales as (  
 -- 월별로 나라별 총 구매 금액 계산  
 select  
 *substr*(o.orderDate, 1, 7) as date\_y\_m,  
 o.country,  
 *sum*(o.total\_amount) as total\_sales  
 from orders\_with\_details o  
 group by date\_y\_m, o.country  
),  
most5\_country\_sales as (  
 -- 월별로 상위 5개 나라를 선정  
 select  
 date\_y\_m,  
 country,  
 total\_sales,  
 *dense\_rank*() over (partition by date\_y\_m order by total\_sales desc) as top  
 from monthly\_country\_sales  
)  
  
-- main query : 상위 5개 나라만 선택  
select date\_y\_m, country, *sum*(total\_sales), top  
from most5\_country\_sales  
where top <= 5  
group by date\_y\_m, country, top  
order by date\_y\_m, *sum*(total\_sales) desc;

-- 가설1 : 주문 월에 따라 가장 많은 구매를 한 나라들이 달라질 것이다.  
-- select date\_y\_m, country, sum(total\_sales), top  
-- from most5\_country\_sales  
-- where top <= 1  
-- group by date\_y\_m, country, top  
-- order by date\_y\_m, sum(total\_sales) desc;

[output - 월별 주문 top 5 확인]

텍스트, 벼슬이(가) 표시된 사진

자동 생성된 설명

텍스트, 벼슬이(가) 표시된 사진

자동 생성된 설명

[code – 연도별 주문 top 5 확인]

## 2. 연도별로 주문 top 5 확인  
 with orders\_with\_details as (  
 -- orders 테이블과 orderdetails 테이블을 조인하여 필요한 데이터를 준비한다.  
 select  
 o.orderNumber,  
 o.orderDate,  
 o.customerNumber,  
 c.country,  
 (od.quantityOrdered \* od.priceEach) as total\_amount  
 from orders o  
 join orderdetails od on o.orderNumber = od.orderNumber  
 join customers c on o.customerNumber = c.customerNumber  
 where o.status = 'Shipped'  
),  
monthly\_country\_sales as (  
 -- 연도별로 나라별 총 구매 금액 계산  
 select  
 *substr*(o.orderDate, 1, 4) as date\_year,  
 o.country,  
 *sum*(o.total\_amount) as total\_sales  
 from orders\_with\_details o  
 group by date\_year, o.country  
),  
most5\_country\_sales as (  
 -- 연도별로 상위 5개 나라를 선정  
 select  
 date\_year,  
 country,  
 total\_sales,  
 *dense\_rank*() over (partition by date\_year order by total\_sales desc) as top  
 from monthly\_country\_sales  
)  
  
-- main query : 상위 5개 나라만 선택  
select date\_year, country, *sum*(total\_sales), top  
from most5\_country\_sales  
where top <= 5  
group by date\_year, country, top  
order by date\_year, *sum*(total\_sales) desc;

[output – 연도별 주문 top 5 확인]

텍스트, 스크린샷, 메뉴이(가) 표시된 사진

자동 생성된 설명

가설1. 주문 월에 따라 가장 많은 구매를 한 나라들이 달라질 것이다.

Cf) 데이터는 2003년 1월부터 2005년 5월까지 존재한다.

1. 2003년도 1월 ~ 5월 사이에는 다르다.  
   그러나 2004년도 1월에 프랑스가 가장 많이 구매를 한 사실을 제외하면, 2003년도 6월부터 2005년도 1월까지 가장 많이 구매한 나라는  
   미국으로 나오게 된다. 가설이 옳다고 하기에는 가장 많은 구매를 한 나라가 달라지지 않는 경향을 보여줬다. 따라서 해당 가설은 기각할 수 있다.

가설2. 성수기(특정 월)에 구매가 급증할 것이다.스크린샷, 라인, 그래프이(가) 표시된 사진

자동 생성된 설명

1. 2003년도와 2004년도의 10월, 11월에 구매가 급증했음을 보여주었다.  
   주어진 자료 내에서는 가설이 어느 정도 타당하다 생각할 수 있다.

가설3. 연도별 가장 많은 구매를 한 나라의 순위는 크게 달라지지 않을 것이다.

1. 미국은 1위를 계속 차지하며 스페인과 프랑스의 경우 서로 2위와 3위를 번갈아가면서 바뀌는 것을 보여준다. 이 외에 4위와 5위에는 뉴질랜드와 호주만 보이는 경향이 있다. 결론적으로 해당 가설은 어느 정도 타당하다 할 수 있다.