

[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 5 개 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 5 개 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;

-- 가설 1 : 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]

2003-01 Norway	50218.95	1
2003-01 Spain	40206.2	2
2003-01 USA	15718.61	3
2003-01 Germany	10549.01	4
2003-02 Denmark	53959.21	1
2003-02 Italy	52151.81	2
2003-02 USA	22292.62	3
2003-03 USA	53415.29	1
2003-03 Philippines	51001.22	2
2003-03 UK	48425.69	3
2003-03 Sweden	7674.94	4
2003-04 Australia	45864.03	1
2003-04 Singapore	44380.15	2
2003-04 Austria	35826.33	3
2003-04 France	33383.14	4
2003-04 USA	21665.98	5
2003-05 France	82096.57	1
2003-05 Spain	57131.92	2
2003-05 USA	32641.98	3
2003-05 Australia	7565.08	4
2003-06 USA	81910.6	1
2003-06 Spain	36251.03	2
2003-06 UK	29429.14	3
New		
2003-06 Zealand	2880	4
2003-07 USA	94276.97	1
2003-07 France	51572.43	2
New		
2003-07 Zealand	32077.44	3
2003-07 Australia	24013.52	4
2003-08 USA	147412.05	1
2003-08 Finland	29716.86	2
2003-08 Belgium	1128.2	3
2003-09 USA	69308.76	1
2003-09 Australia	51376.05	2
2003-09 Spain	44939.85	3
2003-09 Singapore	38350.15	4
2003-09 Finland	32723.04	5
2003-10 USA	193241.34	1
2003-10 Singapore	67392.85	2
2003-10 Finland	37602.48	3
2003-10 Norway	36798.88	4
2003-10 Denmark	36164.46	5
2003-11 USA	290655.54	1
2003-11 Spain	130137.12	2
2003-11 France	116314.25	3
2003-11 Australia	97989.35	4
2003-11 Norway	85024.46	5
2003-12 USA	149660.85	1
2003-12 Spain	53121.69	2
2003-12 Canada	36527.61	3
New		
2003-12 Zealand	20220.04	4
2003-12 Hong Kong	17193.06	5

2004-01 France	98779.88	1
2004-01 Spain	81993.16	2
2004-01 USA	57124.16	3
2004-01 Japan	47177.59	4
2004-01 UK	7310.42	5
2004-02 USA	92948.52	1
2004-02 Switzerland	47375.92	2
2004-02 Australia	44894.74	3
2004-02 Ireland	32538.74	4
2004-02 France	24757.31	5
2004-03 USA	70903.2	1
2004-03 France	40978.53	2
2004-03 Germany	33820.62	3
New		
2004-03 Zealand	31670.37	4
2004-03 UK	24995.61	5
2004-04 USA	36458.26	1
2004-04 France	36069.26	2
2004-04 Canada	29284.42	3
2004-04 Denmark	28211.7	4
2004-04 Spain	26155.91	5
2004-05 USA	117983.85	1
2004-05 Spain	35420.74	2
2004-05 Finland	28394.54	3
2004-05 France	25080.96	4
2004-06 USA	77271.78	1
2004-06 Singapore	44160.92	2
New		
2004-06 Zealand	37281.36	3
2004-06 Canada	22997.45	4
2004-06 Japan	22037.91	5
2004-07 USA	93991.54	1
2004-07 Italy	51619.02	2
2004-07 France	47924.19	3
2004-07 Belgium	45352.47	4
2004-07 Australia	45221.86	5
2004-08 USA	215256.19	1
2004-08 Switzerland	61402	2
2004-08 Italy	48298.99	3
2004-08 Canada	37527.58	4
2004-08 Norway	32260.16	5
2004-09 USA	60664.69	1
2004-09 Sweden	48809.9	2
2004-09 Singapore	38785.48	3
2004-09 Finland	34341.08	4
2004-09 Italy	33924.24	5
2004-10 USA	168723.81	1
2004-10 France	72618.81	2
2004-10 Germany	61234.67	3
2004-10 Denmark	53745.34	4
2004-10 UK	52825.29	5
2004-11 USA	377124.68	1
2004-11 UK	93947.25	2
2004-11 France	88878.94	3
2004-11 Australia	82261.22	4
2004-11 Spain	58090.77	5
2004-12 USA	116603.76	1
2004-12 Spain	116208.4	2
New		
2004-12 Zealand	91606.59	3
2004-12 France	58912.24	4
2004-12 Australia	31835.36	5
2005-01 USA	98334.59	1
2005-01 Finland	92555.86	2
2005-01 Japan	33967.73	3
2005-01 Australia	27083.78	4
2005-01 Belgium	14379.9	5
2005-02 Spain	120166.58	1
2005-02 USA	102466.71	2
2005-02 France	49523.67	3
2005-02 Finland	23602.9	4
2005-02 Australia	21432.31	5
2005-03 USA	113891.81	1
2005-03 France	77017.35	2
New		
2005-03 Zealand	33593.32	3
2005-03 Australia	29848.52	4
2005-03 Sweden	27966.54	5
New		
2005-04 Zealand	108122.9	1
2005-04 France	47348.6	2
2005-04 UK	37258.94	3
2005-04 USA	31755.34	4
2005-04 Singapore	2326.18	5
2005-05 Austria	52420.07	1
2005-05 Spain	46895.48	2
2005-05 Italy	35362.26	3

[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 5 개 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 5 개 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]

	date_year	country	`sum(total_sales)`	top
1	2003	USA	1172200.59	1
2	2003	Spain	369488.73	2
3	2003	France	283366.39	3
4	2003	Australia	226808.03	4
5	2003	UK	161494.57	5
6	2004	USA	1485054.44	1
7	2004	France	506660.01	2
8	2004	Spain	392816.48	3
9	2004	Australia	204213.18	4
10	2004	New Zealand	195592.89	5
11	2005	USA	374949.23	1
12	2005	Spain	185164.80	2
13	2005	France	175724.18	3
14	2005	New Zealand	165343.66	4
15	2005	Finland	116158.76	5

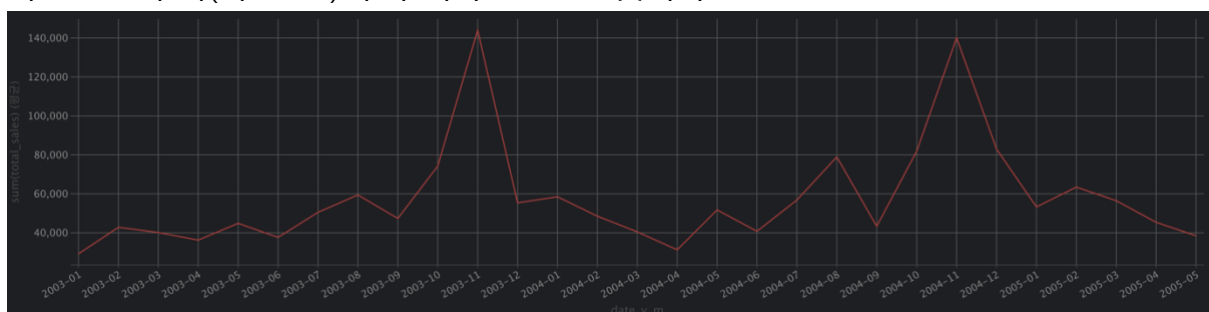
가설1. 주문 월에 따라 가장 많은 구매를 한 나라들이 달라질 것이다.

Cf) 데이터는 2003년 1월부터 2005년 5월까지 존재한다.

a. 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.



a. 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.

- a. 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.