

## Red

You have a table of sales for your fruit cart that looks something like this:

Table: fct\_sales

dateid (date)

fruit (varchar)

sold (int)

Sample table:

| dateid     | fruit  | sold |
|------------|--------|------|
| 2015-01-01 | Apple  | 31   |
| 2015-01-01 | Orange | 19   |
| 2015-01-02 | Apple  | 37   |
| 2015-01-02 | Orange | 26   |
| 2015-01-03 | Apple  | 21   |
| 2015-01-03 | Orange | 21   |
| 2015-01-04 | Apple  | 35   |
| 2015-01-04 | Orange | 27   |

Assume that the table is well-behaved, and every date has a row for apples and a row for oranges.

**Q1: Find the difference in the number of apples and oranges sold each day (apples minus oranges)**

This can be done several different ways. The clunky-but-natural is to construct a table for apple sales and a table for orange sales, then to join them together. The end result would look something like this:

```
SELECT a.dateid, pos_sales - neg_sales
from (
  SELECT dateid, pos_sales
  from fct_sales
  where fruit = 'Apple'
) a
join (
  SELECT dateid, neg_sales
  from fct_sales
  where fruit = 'Orange'
) o
```

on a.dateid = o.dateid

**Q2: Find the difference in the number of apples and oranges sold each day (apples minus oranges), without using a join**

```
select dateid,  
SUM(CASE WHEN fruit = 'Apple' THEN sold WHEN fruit = 'Orange' THEN -sold END) AS diff  
from fct_sales  
group by dateid
```

There are multiple solutions, but this is probably the most elegant. All non-contrived solutions will involve at least one case statement.

**Q3: Suppose that we add multiple types of fruit: bannanas, lemons, plums, etc. Find the total number of fruits sold per day, but only including those fruits which sold at least 10 on January 1st, 2015.**

We can use a self-join, but a slightly more elegant solution is to use a nested subquery. First, we have to get a list of fruits which sold at least 10 on January 1st, 2015:

```
select distinct fruit  
from fct_sales  
where dateid = '2015-01-01'  
and sold >= 10
```

Then, we write a query to get the total amount of fruit sold per day and stick the latter inside the former:

```
select dateid, sum(sale)  
from fct_sales  
where fruit in (  
    select distinct fruit  
    from fct_sales  
    where dateid = '2015-01-01'  
    and sold >= 10 )  
group by 1
```

Notice that this subquery does not need a unique table identifier because the table created can't be referred to other than in the where clause that created it.

We can also do this with joins:

```
select dateid, sum(sale)
```

```
from fct_sales fs
join (
select distinct fruit
from fct_sales
where dateid = '2015-01-01'
and sold >= 10
) xxx
on fs.fruit = xxx.fruit
group by 1
```

Note that the distinct is not strictly necessary but is good practice because a table with two rows saying 'Plum' will double-count plums and fail silently.