Lab 4: OLAP Queries for Filtering

# 1. Overview

This lab continues to examine how to use SQL statements to query a data mart. We will use the approach of “learning with examples”. That is, you will be given a range of examples of SQL queries, and you will need to read them and understand their meanings before running them with a computer.

We will continue to use the data mart file from the last lab. If you have lost this file, please refer to Lab 3 to retrieve the file.

# 2. Filter by Several Dimensions

Firstly, change the mode of SQLite to table by running .mode table. Then, read the following queries to understand why they are written in such ways. Finally, execute the following queries in the SQLite Shell.

Query 4.1a: Show the sales statistics by product category for the sales in USA in 2010.

SELECT P.category, SUM(F.receipts) as receipts, SUM(F.quantity) as quantity,

SUM(F.cost) as cost

FROM Sales F, Product P, Date D, Store S

WHERE F."keyP" = P."keyP" AND F."keyD" = D."keyD" AND F."keyS" = S."keyS"

AND D.year = '2010' AND S.country = 'USA'

GROUP BY P.category;

Query 4.1b: Show the sales statistics by product category for the sales in the states CA and WA in Jan-Feb 2010.

SELECT P.category, SUM(F.receipts) as receipts, SUM(F.quantity) as quantity,

SUM(F.cost) as cost

FROM Sales F, Product P, Date D, Store S

WHERE F."keyP" = P."keyP" AND F."keyD" = D."keyD" AND F."keyS" = S."keyS"

AND (D.month = '2010 Jan' OR D.month = '2010 Feb')

AND (S.state = 'CA' OR S.state = 'WA')

GROUP BY P.category;

Query 4.1c: Show the sales of the brand 'Better'. Break down the numbers by customers' occupation.

SELECT C.occupation, SUM(F.receipts) as receipts, SUM(F.quantity) as quantity,

SUM(F.cost) as cost

FROM Sales F, Product P, Customer C

WHERE F."keyP" = P."keyP" AND F."keyC" = C."keyC"

AND P.brand = 'Better'

GROUP BY C.occupation;

Query 4.2a: Show the sales receipts by product category and by store country.

SELECT P.category, S.country, SUM(F.receipts) as receipts

FROM Sales F, Product P, Store S

WHERE F."keyP" = P."keyP" AND F."keyS" = S."keyS"

GROUP BY P.category, S.country;

Query 4.2b: Show the sales receipts in 2011 by product category and by store country.

SELECT P.category, S.country, SUM(F.receipts) as receipts

FROM Sales F, Product P, Store S, Date D

WHERE F."keyP" = P."keyP" AND F."keyS" = S."keyS" AND F."keyD" = D."keyD"

AND D.year = '2011'

GROUP BY P.category, S.country;

Query 4.2c: Show the monthly sales quantity of products of type 'Pancakes' and 'Coffee' to male customers in stores in cities 'Los Angeles' and 'Beverly Hills' in Jul-Dec of 2010. Break down the numbers by product type and months.

SELECT P.type, D.month, C.gender, S.city, SUM(F.quantity) as quantity

FROM Sales F, Product P, Store S, Date D, Customer C

WHERE F."keyP" = P."keyP" AND F."keyS" = S."keyS" AND F."keyD" = D."keyD" AND F."keyC" = C."keyC"

AND (P.type = 'Pancakes' OR P.type = 'Coffee')

AND C.gender = 'M'

AND (S.city = 'Los Angeles' OR S.city = 'Beverly Hills')

AND D.month in ('2010 Jul', '2010 Aug', '2010 Sep', '2010 Oct', '2010 Nov', '2010 Dec')

GROUP BY P.type, D.month, C.gender, S.city;

# 3. Top N

Query 4.3a: List the sales amount by product type in the top 4 cities in sales receipts in 2011.

SELECT S.city, P.type, SUM(F.receipts)

FROM SALES F

JOIN STORE S ON F."keyS" = S."keyS"

JOIN DATE D ON F."keyD" = D."keyD"

JOIN PRODUCT P ON F."keyP" = P."keyP"

JOIN (

SELECT S.city, SUM(F.receipts)

FROM SALES F

JOIN STORE S ON F."keyS" = S."keyS"

JOIN DATE D ON F."keyD" = D."keyD"

WHERE D.year = '2011'

GROUP BY S.city

ORDER BY SUM(F.receipts)

LIMIT 4

) Top4Cities ON S.city = Top4Cities.city

WHERE D.year = '2011'

GROUP BY S.city, P.type;