COMP33111: Tutorial/lab exercise 3

Guide answers for
Part 1: Understanding OLAP functionalities

1. Discuss the differences between OLTP and OLAP.

   Guide: see slides 77-79, Lecture 3.

2. Describe the role of ranking functions (e.g. *rank*(), *dense_rank*()) in SQL.

   The main role is to provide an extended family of aggregate functions that can simplify queries: see slides 38-40, Lecture 3.

3. Explain slice and dice operations in OLAP. Given a fact table with sales data (for example *sales*(*market#, product#, time#, amount*) – see the lecture notes) and relevant dimension tables, write an SQL statement that slices the cube to select sales only in week 2, and dice it by regions.

   For slice and dice operations, see slides 24-26 and 32-34 (Lecture 3).

   Assuming that we have the following dimension tables:

   Market (Market_ID, City, Region)
   Product (Product_ID, Name, Category, Price)
   Time (Time_ID, Week, Month, Quarter)
   Sales (Market_ID, Product_ID, Time_ID, Amount)

   The following query provides a solution:

```
SELECT M.Region, SUM (Amount)
FROM Sales S, Time T, Market M
WHERE T.Time_ID = S.Time_ID AND T.Week = 'Week2'
      AND S.Market_ID = M.Market_ID
GROUP BY M. Region
```

4. Explain ROLLUP and CUBE extensions in SQL. Calculate the following query given the table (sales) below:

```
SELECT    Time, Region, Department, sum(Profit) AS Profit
FROM      sales
GROUP BY ROLLUP(Time, Region, Department)
```

Time	Region	Department	Profit
2000	Central	VideoRental	75,000
2000	Central	VideoSales	74,000
2000	East	VideoRental	89,000
2000	East	VideoSales	115,000
2000	West	VideoRental	87,000
2000	West	VideoSales	86,000
2001	Central	VideoRental	82,000
2001	Central	VideoSales	85,000
2001	East	VideoRental	101,000
2001	East	VideoSales	137,000
2001	West	VideoRental	96,000
2001	West	VideoSales	97,000

Result:

Time	Region	Department	Profit
2000	Central	VideoRental	75,000
2000	Central	VideoSales	74,000
2000	Central	[NULL]	149,000
2000	East	VideoRental	89,000
2000	East	VideoSales	115,000
2000	East	[NULL]	204,000
2000	West	VideoRental	87,000
2000	West	VideoSales	86,000
2000	West	[NULL]	173,000
2000	[NULL]	[NULL]	526,000
2001	Central	VideoRental	82,000
2001	Central	VideoSales	85,000
2001	Central	[NULL]	167,000
2001	East	VideoRental	101,000
2001	East	VideoSales	137,000
2001	East	[NULL]	238,000
2001	West	VideoRental	96,000
2001	West	VideoSales	97,000
2001	West	[NULL]	193,000
2001	[NULL]	[NULL]	598,000
[NULL]	[NULL]	[NULL]	1,124,000

5. Given are the fact table *PropertySale:*

PropertySale(*branchNo, propertyType*, *yearMonth*, *saleAmount*)

and dimension table *Branch*(*branchNo*, *city*), along with the following SQL statement:

```
SELECT    propertyType, yearMonth, city, SUM(saleAmount)
FROM      Branch, PropertySale
WHERE     Branch.branchNo = propertySale.branchNo
      AND propertySale IN ('2007-01', '2007-02')
      AND branch.city IN (Manchester, Edinburgh, Birmingham)
GROUP BY CUBE (propertyType, yearMonth, city).
```

(a) Explain (in plain English) which data would the above query retrieve.

(b) If the above query resulted in the table on the next page, what would be the result of the following query:

```
SELECT      propertyType, yearMonth, city, SUM(saleAmount)
FROM        Branch, PropertySale
WHERE       Branch.branchNo = propertySale.branchNo
      AND   propertySale IN ('2007-01', '2007-02')
      AND   branch.city IN ('Manchester', 'Edinburgh',
Birmingham')
GROUP BY ROLLUP (propertyType, yearMonth, city).
```

propertyType	yearMonth	city	saleAmount
flat	2007-01	Manchester	115432
flat	2007-01	Edinburgh	236573
flat	2007-01	Birmingham	7664
flat	2007-01		359669
flat	2007-02	Manchester	123780
flat	2007-02	Edinburgh	323100
flat	2007-02	Birmingham	8755
flat	2007-02		455635
flat		Manchester	239212
flat		Edinburgh	559673
flat		Birmingham	16419
flat			815304
house	2007-01	Manchester	77987
house	2007-01	Edinburgh	135670
house	2007-01	Birmingham	4765
house	2007-01		218422
house	2007-02	Manchester	76312
house	2007-02	Edinburgh	166503
house	2007-02	Birmingham	4889
house	2007-02		247713
house		Manchester	154308
house		Edinburgh	302173
house		Birmingham	9654
house			466135
	2007-01	Manchester	193419
	2007-01	Edinburgh	372243
	2007-01	Birmingham	12429
	2007-01		578091
	2007-02	Manchester	2001001
	2007-02	Edinburgh	489603
	2007-02	Birmingham	13644
	2007-02		703348
		Manchester	393520
		Edinburgh	861846
		Birmingham	26073
			1281439

Guide answer:
(a) "Explanation" of the CUBE query: retrieve total amount of sale in January and February 2007 in Manchester, Edinburgh and Birmingham, with subtotals for each property type, month and city (including all cross-tabular subtotals).

(b) Result of the ROLLUP query:
(hint: compare the grouping sets in CUBE and ROLLUP for the given set of attributes)

ROLLUP grouping sets:

(propertyType, yearMonth, city),
(propertyType, yearMonth)
(propertyType)
( )

propertyType	yearMonth	city	saleAmount
flat	2007-01	Manchester	115432
flat	2007-01	Edinburgh	236573
flat	2007-01	Birmingham	7664
flat	2007-01		359669
flat	2007-02	Manchester	123780
flat	2007-02	Edinburgh	323100
flat	2007-02	Birmingham	8755
flat	2007-02		455635
flat		Manchester	239212
flat		Edinburgh	559673
flat		Birmingham	16419
flat			815304
house	2007-01	Manchester	77987
house	2007-01	Edinburgh	135670
house	2007-01	Birmingham	4765
house	2007-01		218422
house	2007-02	Manchester	76312
house	2007-02	Edinburgh	166503
house	2007-02	Birmingham	4889
house	2007-02		247713
house		Manchester	154308
house		Edinburgh	302173
house		Birmingham	9654
house			466135
	2007-01	Manchester	193419
	2007-01	Edinburgh	372243
	2007-01	Birmingham	12429
	2007-01		578091
	2007-02	Manchester	2001001
	2007-02	Edinburgh	489603
	2007-02	Birmingham	13644
	2007-02		703348
		Manchester	393520
		Edinburgh	861846
		Birmingham	26073
			1281439