Name:	
Roll No:	
Class/Sem:	TE/V
Experiment No.:	2
Title:	Implementation of Dimension and Fact tables and perform OLAP operations.
Date of Performance:	
Date of Submission:	
Marks:	
Sign of Faculty:	

Aim: Implementation of Dimension and Fact tables and perform OLAP operations.

Objective: OLAP stands for Online Analytical Processing. The objective of OLAP is to analyze information from multiple database systems at the same time. It is based on multidimensional data model and allows the user to query on multi-dimensional data.

Theory:

- Online Analytical Processing Server (OLAP) is based on the multidimensional data model.
- The main aim of OLAP is to provide multidimensional analysis to the underlying data.

Following is the list of OLAP operations:

1. Roll-up
2. Drill-down
3. Slice
4. Dice
5. Pivot (rotate)

Roll-up:

- The roll-up operation (also called the drill-up operation) performs aggregation on a data cube, either by climbing up a concept hierarchy for a dimension or by dimension reduction.
- Figure 2.1 shows the result of a roll-up operation performed on the central cube by climbing up the concept hierarchy for location.
- This hierarchy was defined as the total order "street < city < province or state < country."
- The roll-up operation aggregates the data by ascending the location hierarchy from the level of city to the level of country.
- In other words, rather than grouping the data by city, the resulting cube groups the data by country.

Drill-down:

- Drill-down is the reverse of roll-up. It navigates from less detailed data to more detailed data.
- Drill-down can be realized by either stepping down a concept hierarchy for a dimension or introducing additional dimensions.
- Figure 2.1 shows the result of a drill-down operation performed on the central cube by stepping down a concept hierarchy for time defined as "day < month < quarter < year."
- Drill-down occurs by descending the time hierarchy from the level of quarter to the more detailed level of month.
- The resulting data cube details the total sales per month rather than summarizing them by quarter.

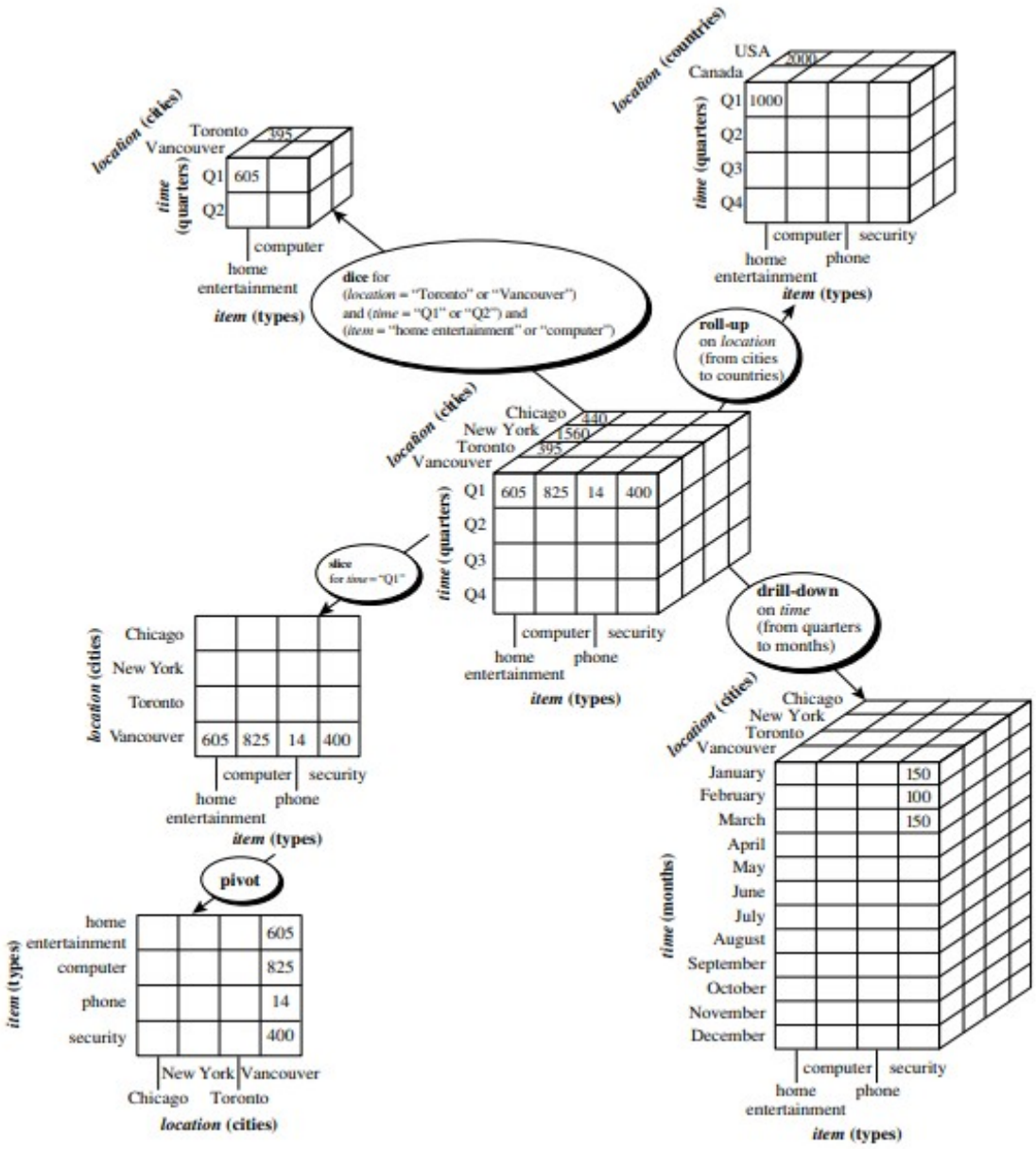

Figure 2.1: Examples of typical OLAP operations on multidimensional data.

Slice:

- The slice operation performs a selection on one dimension of the given cube, resulting in a sub cube.
- Figure 2.1 below shows a slice operation where the sales data are selected from the central cube for the dimension time using the criterion time = "Q1."

Dice:

- The dice operation defines a sub cube by performing a selection on two or more dimensions.
- Figure 2.1 shows a dice operation on the central cube based on the following selection criteria that involve three dimensions: (location = "Toronto" or "Vancouver") and (time = "Q1" or "Q2") and (item = "home entertainment" or "computer").

Pivot:

- Pivot (also called rotate) is a visualization operation that rotates the data axes in view to provide an alternative data presentation.
- Figure 2.1 shows a pivot operation where the item and location axes in a 2-D slice are rotated.

Problem Statement:

To efficiently analyze and visualize sales data, leveraging a multi-dimensional data cube approach. The operations performed will include aggregation, filtering, and reorganization of data to support insightful decision-making.

Code:

```
CREATE TABLE sales_data (
   location VARCHAR(50),
   time VARCHAR(10),
   item VARCHAR(50),
   sales INT
);

-- Example data: Adding a 'country' column to the table for demonstration
ALTER TABLE sales_data ADD COLUMN country VARCHAR(50);

-- Roll-up query: Aggregate sales by 'country'
SELECT
   country,
   SUM(sales) AS total_sales
FROM
   sales_data
GROUP BY
   country;
```

```
-- Drill-down query: Aggregate sales by month (assuming 'time' can be month)
SELECT
   time AS month,
   SUM(sales) AS total_sales
FROM
   sales_data
WHERE
   time IN ('January', 'February', 'March') -- Example months
GROUP BY
   time;


-- Slice query: Select sales data where time = 'Q1'
SELECT
   location,
   time,
   item,
   sales
FROM
   sales_data
WHERE
   time = 'Q1';


-- Dice query: Filter based on location, time, and item
SELECT
   location,
   time,
   item,
   sales
FROM
   sales_data
WHERE
   location IN ('Toronto', 'Vancouver') AND
   time IN ('Q1', 'Q2') AND
   item IN ('home entertainment', 'computer');
```

```
-- Pivot query: Rotate 'item' and 'location' axes
SELECT
  location,
    SUM(CASE WHEN item = 'home entertainment'  THEN sales ELSE 0 END)  AS
home_entertainment_sales,
  SUM(CASE WHEN item = 'computer' THEN sales ELSE 0 END) AS computer_sales
FROM
  sales_data
GROUP BY
  location;
```

Output:

```
INSERT INTO sales_data (location, time, item, sales, country) VALUES
('Toronto', 'Q1', 'home entertainment', 100, 'Canada'),
('Toronto', 'Q1', 'computer', 150, 'Canada'),
('Vancouver', 'Q1', 'home entertainment', 200, 'Canada'),
('Vancouver', 'Q2', 'computer', 250, 'Canada'),
('Toronto', 'Q2', 'home entertainment', 300, 'Canada'),
('Toronto', 'Q2', 'computer', 350, 'Canada');
```

1. Creating the Dimension Tables
2. Creating the Fact Table
3. Inserting values in both dimension and fact tables
4. Displaying the tables
5. Write SQL Queries for all the above OLAP operations.

Conclusion:

Q1. What is the importance of OLAP operations?
Q2. What are the key features of OLAP?